nature

www.nature.com/nature

Vol 445 | Issue no. 7125 | 18 January 2007

Leading by example

The primary safeguard against scientific misconduct is the example set every day by thousands of senior researchers in the laboratory.

Everybody likes a good scandal, and there is nothing like a fresh allegation of research misconduct to set tongues wagging in the scientific community and outside it.

There is broad agreement within the community on two main points regarding outright scientific fraud: it is rare, and it is serious. This unanimity frays at the marginal territory lying just beyond falsification and fabrication of data, however, where 'misconduct' may involve hogging the limelight for one paper too many, or tardy record-keeping at the laboratory bench. But even here, there is a reasonable consensus that scientists in every nation and every discipline will do the best for both themselves and their work by keeping all forms of substandard conduct to a minimum.

It has sometimes been argued that this is so because the credibility of the scientific enterprise is on the line. But if the truth be told, the smattering of egregious misconduct cases that have attracted widespread attention — the most recent being that of South Korea's Woo Suk Hwang — have made little ill impression on the generally favourable public attitude to science as a whole.

It is in fact the reliability of science — or to be more specific, that of the scientific record in the literature — that is the principal victim of scientific misconduct. No aircraft are going to fall out of the sky because somebody fudged a graph in a materials-science journal. Rather, the problem is the time wasted by other scientists chasing up minor and not-so-minor infractions in pursuit of their own research.

A range of tools are available to help minimize research misconduct (see page 240). The most obvious are the systems of investigation and sanction that have been established in many nations. Rather less obvious, but perhaps just as useful, are the methods being developed to secure a scientific environment that is conducive to the highest standards of research conduct.

The thorough and fair investigation of misconduct allegations is a necessary activity that is fraught with difficulty, and likely to remain so. In the United States, the Office of Research Integrity at the health department and the Inspector General's office at the National Science Foundation have between them overseen such investigations at universities for more than a decade now. But every investigation is a new and testing experience for the institution involved.

Parallel developments have occurred in Europe, especially in Germany, where the creation of an ombudsman system at the DFG, Germany's main research agency, in 1999 has provided useful impetus for fair and thorough university-based investigations. Other nations, including Britain, still lack a national mechanism responsible for this — leaving institutions to fend for themselves, with uncertain results.

A more aggressive approach is being taken in east Asia, where major universities have responded sharply to recent misconduct allegations, constituting a major improvement on their past reluctance to act. However, it is not at all clear that this approach will extend beyond high-profile cases that have already found their way into the public domain — or that adequate checks are in place to assure the rights of the accused.

Attention should certainly be paid to nurturing the right framework for investigation and to setting up appropriate sanctions in an effective, cross-institutional manner. But hell will freeze over before an investigatory system can be developed that is universally recognized as comprehensive and fair.

In the interim, the scientific community should continue to concentrate on developing an environment that is inherently consistent with minimizing scientific misconduct.

A respectable level of ethics training for all postgraduate students is an important element of this. It needs to be introduced at all research universities — alongside stricter rules on record-keeping, and arrangements for protecting whistleblowers, where this is missing at the national level.

But most important of all, as the first scientific studies of the factors behind good conduct confirm, is the example set by senior researchers themselves. It is here in the laboratory — not in the law courts or the offices of a university administrator — that the trajectory of research conduct for the twenty-first century is being set. ■

The database revolution

Funding agencies face conflicting challenges in supporting the databases essential to modern biology.

In molecular biology and genetics, high-throughput data gathering has progressed in a short time from a novel strategy of the few to a standard approach of the many.

Much of these data are not collected in a way that guarantees their accessibility to other researchers for future analysis. The best databases, however, are online agglomerations of raw and processed data, analytic tools and insights that have been collectively dubbed 'knowledge environments'.

Funding agencies are slowly adjusting their priorities to reflect the importance of consistent, assured support for some of these databases. The US National Institutes of Health (NIH), for example, has recognized that it must provide guidance for its research sections and grantees so as to maximize the return, in both knowledge and public health, on its investments in biomedical research.

Which strategies best support the collection, analysis and dissemination of large databases of related information? Do knowledge environments promote the aims of the funding agencies? At a meeting in Bethesda, Maryland, last month, it was clear that the NIH is

©2007 Nature Publishing Group

struggling to find a middle road between two diametrically opposed approaches to the development of such databases. Top-down pressure by the agency on researchers to use certain software or formats would probably impede their development. But a bottom-up strategy that merely encourages cross-project cooperation, while allowing researchers total freedom to devise their own databases, is bound to be chaotic, does not guarantee cross-compatibility of data, and ultimately reduces the likelihood that their contents will be used to maximally benefit research.

What is clear is that individual labs can no longer make much progress alone. Currently, many researchers feel they are drowning in data. For all they know, a database might contain answers to patient safety issues or glimmers of new therapeutics — but this is being lost through an inability to effectively harvest the data already available. Other opportunities are missed because both experts and data are 'silo'-ed in isolated and often inaccessible systems. On top of these issues is the fact that neither databases nor the experts that create them are permanent or inseparable.

The NIH and its equivalent agencies elsewhere in the world are now turning their attention to working out how best to assist the growth of validated and accessible databases. This should involve, at the least, development of policies for evaluating proposals on databases and associated analytic tools, for their sustained funding, and for ensuring that the data deposited remain accessible long after the project originators have moved on.

The NIH itself, if it chose, could aim for something grander. It could take it upon itself to define a broad reference model and the basic architecture for knowledge environments. It could even build a centralized warehouse with Google-like storage, a veritable National Biomedical Resource of raw data and the tools to access and analyse them. The challenge then would be how to store, retrieve and read data created on multiple systems by diverse research groups.

But perhaps it will prove more realistic for the US agency to concentrate on improving the inter-operability of databases, rather than pushing for their merger, and to provide incentives for building in 'joins' from the start. The NIH should work on this with industrial companies and other government agencies.

All funders should also be aware of the need to support viable career paths for the software engineers and bioinformaticians who create the knowledge environments and curate the data in them. And in order to obtain value for money, it will be vital for funding agencies to carefully select the databases they choose to support and then to support them for the long term. They must encourage the sustained availability of these data and build incentives for the development of cross-querying capability.

"To obtain value for money, it will be vital for funding agencies to carefully select the databases they choose to support and then to support them for the long term."

As things stands, even popular databases such as PhysioNet (www.physionet.org), which is facing loss of NIH funding after seven years in development, often lack secure funding. In order to get value, the NIH and other funding agencies need to develop policies for such knowledge environments that will allow long-term support of successful ones.

Researchers also need stronger incentives to sustain their own participation in building knowledge environments. At a minimum, contributors should receive a citable acknowledgement of depositions. Leadership and trust are required to ensure that primary researchers personally benefit from storing their data in open databases.

Michael Huerta, associate director of the National Institute of Mental Health, told those at the Bethesda meeting that the NIH is keen to get more mileage out of high-throughput data acquired with public funds. But that will mean that the agency should require the databases it supports to be equipped with tools and protocols that researchers find convenient to use. Ultimately, these databases must show that their data are being used by other groups, and generating work of publishable quality. ■

Capitol calling

A fresh perspective on science and politics.

Scientists typically master the requirements of their vocation early on — delivering a lecture, writing a grant proposal, setting up a lab. And yet even some of the most accomplished scientists remain largely in the dark about the most basic information underpinning their work: the circumstances surrounding the decision to pay them to do the science in the first place.

The world leader in the public funding of science is, of course, the United States. And America's purse-strings are held not by the president, but by the 435 representatives and 100 senators who make up the US Congress. Many scientists have only the vaguest idea about how that much-maligned institution does its business.

On page 248 of this issue, we introduce a new column that is intended to rectify that and to give readers of *Nature* a fuller insight into how and why different topics take their place in America's research and development agenda. Until the end of last year, the column's author, David Goldston, served as chief of staff for the House Committee on Science, helping to set the legislative agenda for science and technology before Congress.

For the past 20 years, Goldston has been at the heart of forging US science and environmental policy. He is now leaving Capitol Hill on the retirement of his long-time boss, Representative Sherwood Boehlert (Republican, New York). In addition to writing exclusively for *Nature*, Goldston will serve as a visiting lecturer in science and environmental policy at the Woodrow Wilson School of Public and International Affairs at Princeton University, New Jersey.

Goldston's column, Party of One, will provide a seasoned insider's take on the interaction between science and politics. From stem cells and energy policy to climate change and space exploration, it will provide a unique perspective on why things work, or don't work, in Washington DC. And stripped of the congressional staffer's customary anonymity, Goldston will be able to offer his own candid opinions on what's really up, on Capitol Hill. As a result, we are confident that our readers in the United States and around the world will gain a valuable insight into some of the forces that help shape science today. ■

©2007 Nature Publishing Group

RESEARCH HIGHLIGHTS

Flower finds its roots

Science doi:10.1126/science.1135260 (2007)

The plant with the world's largest flower (pictured), which spans up to a metre across and stinks of rotting flesh, has found a place in the tree of life almost 200 years after it was first described.

Rafflesia arnoldii is one of a group of parasitic plants termed Rafflesiaceae. With no proper roots, leaves or stems, plants of this group have proved tricky to classify. Even the chloroplast DNA often used to identify plant relationships has failed to provide clues, as many of the marker genes are missing or truncated.

To get round this, Charles Davis of Harvard University Herbaria in Cambridge, Massachusetts, and his team analysed about 11,500 base pairs of mitochondrial DNA from more than 100 species. They show that, surprisingly, the Rafflesiaceae are nestled within the Euphorbiaceae family, in which most species have tiny flowers.

J. HOLDEN

OPTICS

Most sensitive molecule

Optics Lett. **32,** 59–61 (2007)

Researchers have designed an organic molecule that is more sensitive to light than any seen before. Its strong optical response could be put to use in telecommunication systems, for example to build optical switches.

Javier Perez-Moreno of the University of Leuven in Belgium and his colleagues found that linking two different carbon-ring structures through a nitrogen double-bond enhanced the movement of electrons in the molecule when it was struck by light. The measure of this movement, known as the molecule's 'intrinsic hyperpolarizability' was about 50% higher for the new molecule than for the previous record-holder. However, its value still falls short of the theoretical limit by a factor of around 20, leaving chemists with work to do.

IMMUNOLOGY

Opposing forces

J. Exp. Med. doi:10.1084/jem.20061738 (2007)

Researchers have found a protein that safeguards the central nervous system from autoimmune responses.

The protein, known as IL-25, belongs to a family of proteins that regulate inflammation. Daniel Cua and his colleagues at Schering-Plough Biopharma in Palo Alto, California, created mice that lacked IL-25. These mice were more susceptible to an inflammatory autoimmune disease afflicting the central nervous system, suggesting that IL-25 acts against inflammation.

The findings were a surprise because a closely related protein, known as IL-17, promotes inflammation. However, mice that lacked both proteins did not contract the autoimmune disease, demonstrating that the proteins have opposing roles. The researchers believe that IL-25 acts in other organs as well, and could be a target of future therapies for inflammatory diseases.

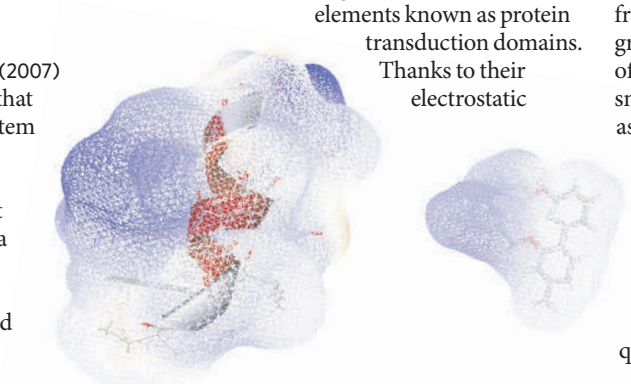
BIOTECHNOLOGY

Delivery service

Nature Methods doi:10.1038/nmeth997 (2007)

The graphic below shows the electrostatic potential of a helical peptide (left) and that of a smaller molecule (right) designed to mimic it: blue and red regions are polar, grey regions non-polar.

The peptide belongs to a class of structural elements known as protein transduction domains. Thanks to their electrostatic properties, such peptides can penetrate the membranes of cells. Biologists take advantage of this, coupling other molecules such as dyes or drugs to the peptides for delivery into cells. The mimic designed by Gareth Williams and David Selwood of University College London, UK, and their colleagues can also carry molecules through the cell membrane. It should be cheaper to synthesize than a peptide and easier to optimize for different cargoes.



ECOLOGY

Invasion of the wild radish

Proc. Natl Acad. Sci. USA **104,** 549–552 (2007)

For invading forces, there is strength in numbers. That this is also the case for many plants is confirmed by a study of an invasive colonizer, the California wild radish (*Raphanus sativus*).

Ecologists led by Norman Ellstrand of the University of California, Riverside, showed that larger, more genetically diverse plots of the widespread weed produce more seeds and fruit per plant than smaller, more interrelated groups. This conforms to the predictions of the 'Allee effect', whereby individuals in a small group are expected to decline in fitness as a result of the scarcity of mates.

QUANTUM PHYSICS

Dead and alive

Nature Phys. doi:10.1038/nphys507 (2007)

In the famous thought experiment, Schrödinger's cat was constrained by the quirks of quantum physics to being both

©2007 Nature Publishing Group

half-dead and half-alive at the same time. In practice, the principle of superposition doesn't extend to such large objects, but physicists have succeeded in making hypothetical cats from small groups of photons or atoms.

Now Chao-Yang Lu and Jian-Wei Pan of the University of Science and Technology of China in Hefei report the biggest photonic cat yet, made by entangling the properties of six photons. This six-photon state is technically known as a Greenberger–Horne–Zeilinger state. The team also produced another kind of six-photon grouping, known as a cluster state. Each will provide new opportunities for testing ideas in quantum communication and computing.

CANCER BIOLOGY

Wrong number

Cancer cell **11,** 25–36 (2007)

A study of the role of aneuploidy — an abnormal number of chromosomes — in cancer has reached the surprising conclusion that aneuploidy can inhibit tumorigenesis.

Many tumour cells have an abnormal chromosome number, but the question of whether this property plays an active role in tumour initiation or progression has remained under debate. Don Cleveland of the University of California, San Diego, and his colleagues found that reduction in the motor protein CENP-E, which helps to separate chromosomes during cell division, generated aneuploidy in mice. In these mice, the rates of some types of cancer did increase, but liver tumours were diminished in size. Aneuploidy also delayed tumour onset in mice lacking a tumour suppressor gene.

CELL BIOLOGY

Rapid response

Mol. Cell **24,** 853–865 (2006)

Noise may help cells to cope with stress, say James Collins of Boston University, Massachusetts, and his colleagues.

Working in yeast, they tinkered with the promoter region of a gene to vary how strongly the gene was transcribed. This changed the amount of noise in the expression of the gene. When the cells were subjected to a transient environmental stress, such as an antibiotic, populations with more variable levels of gene expression tended to survive better than those with steadier expression levels. The team suggests that having more variation makes it more likely that some of the cells will find a satisfactory response to the stress. Some genes may have evolved to be prone to noisy expression, they speculate.

BIOPHYSICS

Water in slow motion

Proc. Natl Acad. Sci. USA doi:10.1073/pnas.0601639104 (2006)

Water is mysteriously sluggish inside cells of the extremely halophilic Dead Sea archaea *Haloarcula marismortui*, report Giuseppe Zaccai of the Laue-Langevin Institute in Grenoble, France, and his colleagues.

Using neutron scattering, they found that about three-quarters of the intracellular water in these organisms has a diffusion coefficient 250 times lower than that of bulk water. The idea that cell water is 'different' from bulk water is old and controversial, but it seems that in this salt-loving extremophile it may be true. Why? The researchers aren't yet sure, but suspect that proteins and high concentrations of potassium ions in the cells may somehow conspire to alter the water's liquid-state structure.

M. DURHAM/MINDEN PICTURES/FLPA

NEUROSCIENCE

Bats' brains differ to rats'

Nature Neurosci. doi:10.1038/nn1829 (2007)

A study of bat brains casts doubt on the prevailing theory of how the hippocampus, a brain region concerned with navigation and memory, does its job.

Researchers had based their models of how hippocampal cells function in mammals on observations of rats. But Nachum Ulanovsky and Cynthia Moss at the University of Maryland in College Park report that the brain waves generated by hippocampal cells are different in bats to in rats.

In freely roaming rodents, the hippocampal brain waves, or theta rhythms, are continuous and are assumed to be essential for remembering spatial information. But in bats using echolocation, rather than movement, to explore their surroundings, the theta rhythms appeared in short intermittent bouts. The researchers suggest that theta-wave behaviour may be tailored to a species' mode of navigation.

JOURNAL CLUB

Immanuel Bloch
Johannes Gutenberg University, Mainz, Germany

A cold-matter physicist is amazed by atoms' ability to divide themselves up equally.

Imagine having a box containing an even number of objects, *N*. You want to divide them into two boxes, each of which contains exactly *N*/2 objects. Sounds easy, right?

But let's complicate things a bit. Let's suppose you can't count the objects, nor look at them. Will you still be able to make the split fairly?

A collaboration of researchers from the Massachusetts Institute of Technology and Harvard University, both in Cambridge, recently showed that it's possible to do so for atoms. They divided into two equal halves a Bose–Einstein condensate of 1 million sodium atoms (G.-B. Jo *et al. Phys. Rev. Lett.*, in the press; preprint at http://arxiv.org/abs/cond-mat/0608585).

Bose–Einstein condensates are a novel state of matter that forms at a temperature close to absolute zero. They behave like quantum entities with pronounced wave-like properties. These are properties that I exploit in my own work with condensates, and they also underpin the atom division.

The Cambridge team stored their matter waves in microfabricated magnetic traps, made out of thin wires. The researchers changed the currents in the wires to split slowly the one potential well that was holding the atoms into two.

In a non-interacting gas, this splitting process would probably give a skewed distribution of atoms, and the distribution would be different every time. In this case, the quantum interactions favour a system in which each well contains exactly *N*/2 atoms.

In fact, the evidence suggests that the splitting is accurate to within 50 atoms. I find that truly remarkable from a fundamental point of view. More practically, this dividing of atoms could also be useful in building novel atom interferometers and atomic clocks.

©2007 Nature Publishing Group

NEWS

Europe moves to secure its future energy supply

A sweeping proposal for a common energy policy for Europe, unveiled last week by the European Commission, calls for a radical shift towards a low-carbon economy. Although environmental campaigners have criticized the proposal for not going far enough, other experts say that the plan is much more ambitious and concrete than any suggested before.

The commission recommends a combination of regulatory and technological measures to secure the future energy supply of the European Union (EU) and to reduce the risk of dangerous climate change. It says that the package would help to cut the EU's greenhouse-gas emissions by at least 20% (compared with 1990 levels) by 2020.

"We can improve our prospects for energy security and carbon emissions if these recommendations are followed through."

The commission suggests that renewable energy sources in the EU should make up 20% of the overall energy mix by 2020. Energy generated from wind, solar and other renewable sources currently account for less than 7% of total energy consumption. The commission also calls for biofuels to provide at least 10% of the energy used by the transport sector by 2020.

Imported fossil fuels are projected to remain the main source of energy, however, and the proposed policy includes a commitment to equip the next generation of coal-fired power plants with the technology to capture and store carbon dioxide, which would enable them to operate with low or zero greenhouse-gas emissions. It is the first time that the use of such technology has been proposed at such a high political level.

The report will be discussed by the EU's heads of state at a summit meeting in March. If approved, the recommendations must be translated into national law in all 27 member states.

Policy analysts doubt that all the proposed actions and targets will be accepted. But given the vagueness of previous reviews, experts say that they appreciate the concreteness and level of detail of the proposed changes. "We can improve our prospects for energy security and carbon emissions if these recommendations are followed through, even though it will be difficult to get agreement on all of them," says John Mitchell, an associate fellow with the energy environment and development programme at London-based Chatham House.

And the energy industry, although critical about some of the details, has welcomed the main thrust of the report. "We do know that our fossil future is limited," says a spokeswoman for the Germany energy company RWE.

The most controversial part of the proposal may be the call for a single energy market within the EU. The report calls for trade barriers to be lifted and for the production of energy to be separated from its distribution, to increase competition and allow citizens to choose to buy 'green' energy if they want to as well as reducing dependence on any one source of energy. In the past, energy companies in Germany and France have blocked such moves, and they are expected to keep up their resistance.

Palaeontology journal will 'fuel black market'

A controversial new 'amateur' palaeontology journal has angered academic researchers, who fear that the project will give a sheen of scientific legitimacy to the dealings of commercial fossil hunters.

The organizers of the *Journal of Paleontological Sciences*, launched this month, say that it will publish details of privately held fossils, bringing them in from the "scientific darkness". But traditional palaeontologists say that the journal undermines the field and could fuel the black market in fossil specimens.

The journal is inviting anyone, including commercial fossil hunters and keen amateur collectors, to publish details of their finds. "Lurking in many private collections are valuable fossil treasures. This is the first journal that gives an outlet to the amateur and commercial palaeontologist to try their hand at writing a scientific article," says Ken Carpenter, a member of the journal's publication committee and curator of vertebrate palaeontology at the Denver Museum of Nature and Science in Colorado.

But academics are upset that the fossils will not be in the public domain, and will therefore be unavailable for future study. "This is absolutely imperative in order for our science to exist," says Mark Goodwin at the University of California Museum of Paleontology in Berkeley.

Palaeontologists who publish their results in traditional journals, such as the *Journal of Vertebrate Paleontology*, agree to abide by a code of ethics, whereby they catalogue their specimens in a recognized repository, most often a national museum. Goodwin thinks that the new journal is an attempt to circumvent this practice.

The fears are fuelled by the fact that many of the fossils that change hands in commercial deals do so on the black market. "Promotion of this online journal will encourage the commercialization of vertebrate palaeontology, which unfortunately results in increased illegal collecting activities on public lands in the United States, and the illegal trade and export of fossils from China, other regions of Asia, and the former Soviet Union," Goodwin argues.

Walter Stein, chair of the journal's publication committee and president of PaleoAdventures in Belle Fourche, South Dakota, disagrees. He thinks that some exposure for the privately held

©2007 Nature Publishing Group

A European report suggests fitting all new coal-fired power stations with carbon-capture technology.

A. FOX/CORBIS

The technology does not yet exist to achieve all the targets suggested by the commission. So the report says that investment in research into energy efficiency and clean energy, including nuclear power, should be increased by 50%. The commission plans to detail suggested research goals and spending in a separate plan to be released later this year. But it does say that up to 12 large-scale demonstration projects of different capture and storage technologies should be built by 2015.

Champions of carbon sequestration have greeted the plan as a vital step towards setting climate-friendly performance standards for use of fossil fuels. But given the time it will take to develop the necessary technology, they urge policymakers to get cracking. "Time is of the essence," says Jonathan Gibbins, an energy engineer at Imperial College London. "The technology must be very strongly proven, which is why you cannot get on top of this too early." To meet the 2020 goal, he says that spending would need to be authorized within a year from now.

PETITION AIMS TO MAINTAIN CHEAP DRUGS
A court case in India threatens to derail generic medicines.
www.nature.com/news

Energy and climate issues will be priorities for the EU in the first half of this year. They will also be on the agenda of the G8 meeting in June, and commission president José Manuel Barroso hopes that the firm commitment by the EU to cut greenhouse-gas emissions will provide a nucleus for an international agreement that is at least as ambitious. The Kyoto Protocol, which requires developed countries to cut their 1990-level emissions by an average of 5.5%, runs out in 2012. The commission believes that a renewed agreement should lead to an overall 30% cut in emissions from developed countries, including the United States, by 2020.

But given its stagnant population and economy, experts say it might be less painful for the EU to significantly cut emissions than for the more dynamic US market. "What the commission wants sounds good for Europe, but you must put yourself in other people's shoes," says Gibbins. "You can't go to international negotiations and demand more pain from others than from yourself."

Last week's commission report also contained a section on the costs of climate change, which argued that global warming must be limited to no more than 2 °C compared with pre-industrial levels if catastrophic consequences of warming are to be avoided. Among the projected consequences of inaction outlined in the report are reduced crop productivity in southern Europe compared with increased productivity in northern Europe, a greater number of heat-related deaths, damage due to flooding and negative impacts on tourism. ■

Quirin Schiermeier, with additional reporting by Lucy Odling-Smee

Fossil hunters can now publish without making their specimens available.

J. BLAIR/CORBIS

specimens is better for science than none at all, even if the fossils themselves remain behind closed doors. He adds that there will always be "ten per cent of academics we will never convince, who think we are only in it for the money".

In an apparent bid to deflect accusations that they are interested mainly in promoting commercialization of the field, the journal's publishers have changed their name from the American Association of Paleontological Suppliers to the Association of Applied Paleontological Sciences. The nine-member publication committee will review submissions before drawing on a 22-member pool of potential peer reviewers. "If you can prove it, we will publish it," Stein says.

Goodwin says that the palaeontologists who have agreed to review submissions "should be ashamed of themselves", adding that their efforts would be better spent encouraging private collectors who hold fossils to consider other means of allowing access and study. Donating fossils to a museum and enlisting the help of an academic researcher would ensure that the fossils can be studied in perpetuity, he argues. "When this occurs, science wins over greed."

Commercial fossil trading in the United States has its roots in the 1960s, when private landowners began to realize the potential value of fossils found on their property, causing academic researchers on tight grants to worry that they were being priced out of the field. Typically, the owner of a property can expect to receive up to 25% of the sale value of a specimen, and some US landowners reportedly earn as much as $25,000 a year from the practice. Many collectors also buy fossils as an investment, similar to dealing in artworks or antiques, which fuels the activities of commercial fossil brokers.

In a bid to confront these activities, the US government offers a tax deduction on individuals or estates who donate specimens to national museums. Many museums offer to make a high-quality cast in exchange for the genuine article, in a bid to encourage collectors to donate their treasures. ■

Michael Hopkin

©2007 Nature Publishing Group

Concern as revived 1918 flu virus kills monkeys

The 1918 influenza virus, which killed some 50 million people worldwide, has proved fatal to macaques infected in a laboratory. The study follows *Nature*'s controversial publication[1] of the virus's sequence in 2005, alongside a paper in *Science* that described the recreation of the virus from a corpse and its potency in mice[2].

Some scientists question the wisdom of reconstructing such a deadly virus. Do the benefits outweigh the risks?

Those who carried out the macaque study say yes, as a better understanding of how it acts in a system similar to humans' will help scientists treat future pandemics. The study was carried out in the biohazard level 4 labs of the Public Health Agency of Canada in Winnipeg. Yoshihiro Kawaoka of the University of Wisconsin-Madison and his colleagues infected macaques with the 1918 virus or a contemporary flu strain[3]. Whereas the contemporary virus caused mild symptoms in the lungs, the 1918 flu spread quickly throughout the respiratory system and the monkeys died within days. The damage parallels reports of human patients in 1918.

"The virus caused the monkeys' immune systems to go into overdrive."

The team reports that the 1918 virus caused the monkeys' immune systems to go into overdrive, causing immune proteins to be expressed at abnormally high levels and attack the body — what immunologists call a cytokine storm.

The research suggests that 1918 flu might work in a similar way to other viruses, such as West Nile, that can also cause a massive autoimmune reaction. This suggests a route towards treatment, says Michael Gale, a virologist at the University of Texas Southwestern Medical Center. Drugs that target over-zealous immune responses, such as those that control an immune protein called interleukin-6, are being developed for other diseases. Tweaked versions might work for pandemic flu.

But despite the promise of treatments, the results echo what had already been found in mice, and Gale feels there is a more important issue to be addressed. "The pathogenesis is interesting," he says. "But the key question is: how was it spread so efficiently?"

A team at the Mount Sinai School of Medicine has already started to investigate. Peter Palese is working with Adolfo Garcia-Sastre and Jeffery Taubenberger, who first reconstructed the virus, to find out how it spreads. Working with ferrets, they have found that a change of only one or two amino acids in the flu sequence is enough to stop transmission. They will publish the result in *Science*. Identifying which sections of the genome are responsible for transmission "has huge predictive value for whether strains will become pandemic or not", says Guus Rimmelzwaan at the World Health Organization's National Influenza Centre in Rotterdam, the Netherlands.

The next move for Kawaoka's team is along similar lines — they will be swapping sections in and out of the virus to establish exactly which bits make it so lethal.

But the latest results haven't assuaged everyone's concerns. Richard Ebright, a bacteriologist at Rutgers University, New Jersey, believes the virus should never have been recreated. "The key implication is that the material is now present in at least two locations," he says. The new study, he argues, increases the risk that the virus could escape and sets "a dangerous precedent" for other labs to follow.

Ebright argues that publishing the study in *Nature*, when similar research on more mundane pathogens regularly appears in lower-impact journals, could in itself increase the proliferation risk, if it tempts research groups to work on high-risk pathogens simply to get more recognition. Similar views were expressed off the record by other scientists. Ritu Dhand, *Nature*'s chief biological sciences editor, defends the decision to publish, arguing that because the 1918 virus is not like other flu viruses, gaining insight into what makes it is so virulent in humans is of scientific interest.

Gale agrees that understanding the 1918 flu strain better could have huge public-health benefits. But he says there might be better ways to study this, and admits that some research might be driven as much by historical interest as by the potential health benefits.

Jens Kuhn, a virologist at Harvard Medical School who advises on arms control, also feels divided. "Everything I say, I make 'enemies' on one or the other side," he says. "I am torn sometimes between the two worlds." ■

Kerri Smith

1. Taubenberger, J. K. *et al. Nature* **437,** 889–893 (2005).
2. Tumpey, T. M. *et al. Science* **310,** 77–80 (2005).
3. Kobasa, D. *et al. Nature* **445,** 319–323 (2007).

ON THE RECORD

"Who wants an ugly, stupid kid? I mean, come on."

Jennalee Ryan, head of 'embryo brokerage' the Abraham Center of Life, which offers test-tube embryos created from attractive and intelligent egg and sperm donors, and which is being investigated by US officials.

SCORECARD

Human drugs
The US Food and Drug Administration gave the green light to just 18 new drugs in 2006, an eight-year low...

Dog drugs
...although it has just approved Slentrol, the first drug to battle canine obesity.

NUMBER CRUNCH

50% is the reduction in greenhouse-gas emissions over the past five years claimed last week by budget airline Ryanair, the self-styled "greenest, cleanest airline".

50% is the emissions reduction that the airline was originally going to claim *per passenger*, as revealed by an earlier leaked version of the same press release.

300% is the actual increase in the airline's overall emissions, owing to vastly increased passenger numbers.

SHOWBIZ NEWS

Christmas visits to the American Museum of Natural History shot up 20% relative to the previous year, after the release of Ben Stiller comedy vehicle *Night at the Museum*.

Sources: AP, Bloomberg, FDA, Science Blog, BBC Newsnight

SIDELINES

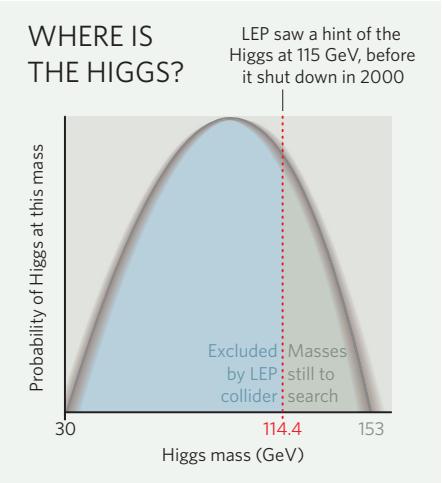
©2007 Nature Publishing Group

Behind the hunt for the Higgs boson

Tension is building on both sides of the Atlantic in the race to find the Higgs boson, the elusive particle predicted by physicists' understanding of mass.

Scientists working at the Tevatron particle collider at Fermilab in Batavia, Illinois, welcomed the new year by measuring the mass of another subatomic particle, the W boson. This finding lowers the expected mass of the Higgs, bringing it further into the range in which the Tevatron might be able to find it.

Meanwhile, the Tevatron's more powerful competitor, the Large Hadron Collider (LHC), is entering its final stages of construction at CERN, the European particle-physics lab near Geneva in Switzerland. If the Higgs exists, the LHC should find it, but the spark of a chance that the Tevatron might get there first is energizing both teams. "We are pulling out all the stops," says Mark Lancaster, one of around 700 scientists working on the Collider Detector at Fermilab (CDF) — an instrument that collects and analyses debris from proton–antiproton collisions in the Tevatron.



Jim Virdee at CERN is the spokesman for one of the LHC's experiments, the Compact Muon Spectrometer. "We are aware of this competition," he says. "It makes us make sure the experiments and the accelerator are ready on time." The enormous parts of the spectrometer are currently being lowered delicately into place in the LHC's tunnels.

The Higgs is thought to be a missing piece of the standard model of particle physics, which lumps together what is known about the fundamental forces and particles. Theories that extend the standard model without the Higgs tend to be more complex than those that include it — leaving physicists hopeful that it will turn up.

Now that estimates of the Higgs' mass have shrunk, the odds have improved that the Tevatron will be able to detect it. In 2004, the upper bound of the boson's mass hovered around 250 gigaelectron volts (GeV), which was beyond the Tevatron's upper limit of about 170 GeV.

But estimates of the Higgs' mass are based on the mass of other particles, such as the top quark and W boson, and updates in these measurements have since dragged the upper limit on the Higgs down. Most recently, on 8 January, the CDF measured the mass of the W boson more precisely than ever before, at 80.413 GeV.

That figure is consistent with previous results, but towards the upper end of the range, and brings the upper limit for the Higgs down to 153 GeV from 166 GeV. The Large Electron–Positron (LEP) collider, which CERN hosted until it was shut down in 2000, ruled out masses of up to 114.4 GeV — leaving just a small window still to explore (see Graphic).

At the same time, the Tevatron has overcome some of the performance shortfalls that had dampened hopes of finding it at any mass. So researchers there, despite the threat of budget cuts forcing a temporary closure, are allowing themselves some optimism. "We have got about three years," says Lancaster. The Tevatron will run until 2009 and, although the LHC is due to switch on for a warm-up run in November, it will not start to collect data seriously until spring 2008.

Ironically, the Tevatron's next Higgs announcement might be about not finding it. Just before LEP closed to make way for the LHC, it saw hints of a Higgs at around 115 GeV. The thought that the Higgs had narrowly escaped their grasp has haunted CERN's physicists ever since, but the Tevatron is on the verge of having enough data to rule out a particle of that mass. Unless something emerges in the next few months, the LEP ghost will be laid to rest. "I would say that we would be able to rule out this hint from LEP before the LHC gets going in earnest," says Lancaster. ■

Jenny Hogan

©2007 Nature Publishing Group

PURDUE UNIV.

Misconduct? It's all academic...

The legal quagmire, strain and bad press of misconduct investigations leave many universities tempted to ignore misconduct allegations. But getting an investigation right can reduce the pain and boost an institution's reputation, says **Geoff Brumfiel**.

Rumours were flying at Purdue University nuclear engineering school in West Lafayette, Indiana, late in 2005. A recently hired physicist, Rusi Taleyarkhan, claimed to be pioneering a new form of energy called bubble fusion. But researchers both in and outside Purdue had concerns. Questions ranged from whether the neutron radiation in his experiment was from fusion or another source, to whether a supposedly independent confirmation of his findings was ghostwritten by Taleyarkhan.

Lefteri Tsoukalas, then dean of the school, asked the administration to launch a misconduct investigation. What happened next Tsoukalas describes as "a real eye-opener". For a month, he says, the administration ignored his request. An inquiry began only in March 2006, after *Nature* reported his and others' concerns (see *Nature* **440**, 132; 2006).

But even since then, Tsoukalas says, little has been done. In October 2006, he resigned (see *Nature* **444**, 664; 2006), saying that he has "little faith that Purdue is capable of running a research integrity investigation". Taleyarkhan has denied all wrong-doing. Purdue declines to comment.

The situation benefits no one, but is far from unique. Research ethicists and university misconduct officers say that academia's handling of scientific misconduct is often capricious and incomplete. Reasons are complex, but essentially come down to three issues: investigations are laborious; they require legal and scientific expertise that administrators often don't have; and universities may fear that a guilty verdict will stain their reputation.

Countries vary in their approaches, but a look at US universities highlights the problem. Many ethicists told *Nature* the system is riddled with inconsistencies. "I wouldn't even go so far as to glorify the current arrangement as 'a system'," says Arthur Caplan, a bioethicist at the University of Pennsylvania in Philadelphia.

Little guidance

Most US misconduct proceedings follow a similar arc. If an accusation is received, administrators launch a preliminary inquiry. If that finds enough evidence, the university can open a full investigation. But beyond that, there's little guidance. Organizations such as the National Academies and the government's Office of Research Integrity (ORI) provide guidelines on research conduct, but no single document lays out investigative procedures.

Trouble can begin before the inquiry, says Kristina Gunsalus, a lawyer at the University of Illinois at Urbana-Champaign, who has led many proceedings. It's often unclear where whistleblowers should take their concerns, so young students end up discussing them with senior colleagues. The gossip that follows can damage the accused and prejudice an inquiry.

Whistleblowers can also be the target of reprisals. Young people in particular are rarely protected, according to Barbara Redman, dean of nursing at Wayne State University in Detroit, Michigan, who has studied the issue. "Policies are incomplete," she says. Many lose their jobs, regardless of the outcome of their allegations.

Once an investigation starts, it can quickly enter legal limbo. Misconduct is not a legal charge, so in one sense it comes down to university policy. But if anyone is defamed, the university can find itself with a libel case. So it has a strong incentive to be secretive.

Then there's the pressure on investigators, generally scientists at the university, who are expected to serve as judge and jury to their colleagues. This can be awkward. Proceedings are often protracted and complex, with months of hearing testimony and evidence. "Every committee I have ever worked with comes back with these enormous circles under their eyes," says Gunsalus. "I cannot overstate the costs of these procedures for the people involved."

In part because the panels are made up of working scientists, they can deliver unexpected results. One recent example is the University of Pittsburgh's investigation into Gerald Schatten, an obstetrics professor. Schatten collaborated heavily with disgraced stem-cell scientist Woo Suk Hwang of Seoul National University (SNU) in South Korea. The SNU's investigation found that Hwang had faked many of his acclaimed cloning results, and a panel at Pittsburgh was

©2007 Nature Publishing Group

J. FEINGERSH/ZEFA/CORBIS

HOT TOPICS
Join the discussion about controversial stories on *Nature*'s Newsblog.
http://blogs.nature.com/news/blog

charged with determining whether Schatten was also guilty of misconduct.

The panel concluded he was not, but found him guilty of 'research misbehaviour' — a term undefined by university policy and not included in the original charges. "What does it mean and what are the consequences?" asks Gunsalus. "I wish they had done better."

It's not surprising that the legal quagmire, strain and bad press leave many universities tempted to ignore allegations. "Universities have an incentive not to pursue them," says Redman. "They don't want to disgrace their faculty," adds Caplan. "It hurts morale."

Similar problems exist elsewhere. In Britain, for example, major funding bodies such as the research councils and the Wellcome Trust only award grants to universities that have established systems for dealing with misconduct allegations. But there is no body with the power to oversee investigations when they occur.

What can be done? One possible solution is to take investigations out of universities' hands and pass them to an independent body. But this can be controversial, as scientists don't take kindly to investigation by outsiders. The ORI briefly tried to take the lead in the late 1980s, but its heavy-handedness and lack of scientific expertise led to a backlash, according to Caplan: "The ORI blundered around like a brontosaurus in the research misconduct park."

Today, the ORI takes a lower-key role, providing guidelines and reviewing cases involving federal funds. Despite the failings of universities, most people now agree that in a country such as the United States, with a highly privatized university system, it's best to handle misconduct investigations on campus. "There are no better alternatives," says Gunsalus.

Scandinavian countries seem to fare better with independent supervision. In Denmark, whistleblowers can turn to a government body called the Danish Committees on Scientific Dishonesty (DCSD) if they think university boards have not investigated a claim properly. Finland and Norway have similar committees.

Although some of the DCSD's decisions have been controversial, such as when it criticized climate sceptic Bjørn Lomborg for cherry-picking data in his 2001 book *The Skeptical Environmentalist*, it is generally seen as a model for small countries in which it is hard to find independent investigators at the local level.

However, even DCSD member Vagn Lundsgaard Hansen, a mathematician at the Technical University of Denmark in Lyngby, believes claims should be dealt with by the university concerned where possible.

"Asian universities are now showing the world how to conduct fast, fair and transparent investigations."

One way to help universities is to establish common systems and guidelines — although persuading them to comply isn't easy. After Germany's main funding agency, the DFG, asked universities to adopt guidelines in the late 1990s, it finally had to threaten to withdraw eligibility for funding if they didn't comply by mid-2002. The guidelines include establishment of university and DFG ombudsmen to whom whistleblowers can turn with concerns about a lab, and a committee that adjudicates on cases forwarded to it by the ombudsmen.

The universities reluctantly adopted the guidelines. But the ombudsmen expressed concern last October that universities are still not taking referred cases seriously. They complained that university presidents and rectors play down cases, particularly concerns brought by young scientists against more senior colleagues. And when investigations do get started they drag on. The ombudsmen pledged to increase pressure on universities to improve — perhaps by requiring DFG grant application forms to refer to the university's guidelines, thus forcing department heads, who co-sign applications, to explicitly endorse them.

The key may be to convince universities that getting it right can bolster their reputation. The Massachusetts Institute of Technology (MIT) in Cambridge was widely praised for its investigation of immunologist Luk van Parijs, begun in August 2004. MIT moved fast to determine if an investigation was needed, handled complaints confidentially, found him guilty of fabricating results and fired him in October 2005.

Some Asian universities — once notorious for brushing misconduct claims under the carpet — are now showing the world how to conduct fast, fair and transparent investigations.

Quick justice

Last August members of a lab run by veteran biologist Akio Sugino at Osaka University claimed he had faked data in two papers. Two days later the biosciences school to which Sugino belonged formed an investigation committee, and released its conclusion that Sugino had fabricated data the next month. The 40-page report was published on the school's website and in December Sugino was dismissed.

That month the University of Tokyo concluded its investigation, begun in April 2005, into RNA researcher Kazunari Taira and his assistant. It reported that they were unable to provide notebooks to back up their results or to reproduce their experiments and, although it stopped short of finding fraud, fired both.

And in South Korea, the SNU completed its investigation into Hwang in four weeks, publishing its 50-page report on 10 January 2006.

Hisato Kondoh, dean of Osaka University's biosciences school, describes the secret of a successful investigation. "It's important to judge things based on facts, and make the results open as soon as possible, while protecting the rights of people who both reported allegations and are suspected," he says. "If the investigation takes time, everybody gets exhausted. Settling the case at the earliest possible opportunity and bringing the situation back to normal benefits all of us." ■

With additional reporting by Alison Abbott, David Cyranoski, Ichiko Fuyuno, Jim Giles and Lucy Odling-Smee.

See Editorial, page 229.

AHN YOUNG-JOON/AP

Different approaches: Purdue University (top left) is accused of dragging its feet, whereas the investigation of Woo Suk Hwang (above) in South Korea started promptly and took just four weeks.

©2007 Nature Publishing Group

Breeding cheats

Understanding the social and psychological factors behind scientific misconduct will enable bad practice to be minimized, but never eliminated, says **Jim Giles**.

Take one prestigious laboratory. Add some pressing grant deadlines and a dash of apprehension about whether the applications will succeed. Throw in an overworked lab head, a gang of competitive postdocs and some shoddy record-keeping. Finally, insert a cynical scientist with a feeling that he or she is owed glory. It sounds hellish, but elements of this workplace will be familiar to many researchers. And that's worrying, as such an environment is, according to sociologists, the most fertile breeding ground for research misconduct.

Just a decade ago, such a statement would have been speculation. But sociologists are increasingly confident that they understand why scientists cheat. Studies of disgraced researchers, a series of high-profile misconduct cases, and a stream of government funding have created the discipline of research into research integrity. The results are a better understanding of those who betray science, and of the climate in which they do so. They also suggest how misconduct can be reduced, although there are good reasons to think it will never be eliminated.

To tease apart the factors behind acts such as fabricating data or unfairly appropriating ideas, sociologists say we must turn away from the media glare that surrounds extreme cases such as Woo Suk Hwang, the South Korean stem-cell scientist who faked high-profile papers on human cloning[1]. Such cases are to research integrity what serial killers are to crime prevention, says Kenneth Pimple, an ethicist at Indiana University in Bloomington. Hwang and others grab the headlines, but minor acts of misconduct are much more common, and potentially more damaging to scientific progress.

That runs against the grain of traditional thinking on misconduct, at least among scientific societies, which have often argued that cheating is due mainly to a few bad apples. But that view now looks much less tenable. When scientists funded by the US National Institutes of Health were asked in 2002 about misconduct, a third said they had committed at least one of ten serious errant acts, such as falsifying data or ignoring important aspects of the regulations regarding human subjects[2]. Young physicists also returned disturbing results when questioned by the American Physical Society in 2003: more than 10% had observed others giving less than truthful descriptions of research techniques or analyses, for example.

To understand what is driving these figures, researchers would like to study confirmed cases of misconduct. Information here is sparse, as convicted scientists do not generally rush to tell their stories.

What data there are lie mainly in the files of the US Office of Research Integrity (ORI), which oversees biomedical misconduct investigations in the United States. Mark Davis, a criminologist at Kent State University in Ohio, trawled 92 ORI cases from 2000 and earlier and revealed seven factors frequently associated with misconduct[3]. Some involve research climate, such as a lack of support from superiors or competition for promotion. Others, such as a tendency to blame the difficulty of a particular experimental task, point to ways in which individuals justify their own actions.

"Employees are more likely to behave unethically if they believe their managers are treating them unfairly."

Habit-forming

To get a finer-grained image of these factors, sociologists turn to survey data. At the University of Oklahoma in Norman, for example, researchers asked doctoral students how they would react to specific ethical dilemmas, such as if a researcher takes an idea from a colleague and, without letting that colleague know, uses it to apply for funding.

Many of the factors seen in Davis's survey cropped up again, but by asking students about their past and present experiences the Oklahoma team, which presented its work at the ORI conference in Tampa, Florida, last year, added new details. In particular, its work suggests that past experience, such as graduate training, can be more important than the current climate in which people work. Day-to-day aspects matter — interpersonal conflict was associated with unethical decisions, for example — but former experiences, such as having worked in a lab where the head showed positive leadership, seem to be more important.

The need for support extends beyond the level of research groups. Brian Martinson from the HealthPartners Research Foundation in Minneapolis and his colleagues have studied misconduct using the theory of organizational justice, which states that employees are more likely to behave unethically if they believe their managers are treating them unfairly. Sure enough, Martinson's survey respondents were more likely to admit misconduct if they felt that governing structures, such as funding review bodies, had treated them badly[4]. In an as-yet unpublished role-playing exercise, Patricia Keith-Spiegel of Simmons College in Boston, Massachusetts, also found that researchers were more likely to act unethically if a negative decision by a review board was not properly explained.

With these results in place, misconduct experts can make tentative statements about how to limit problems. Robust and positive mentoring is top of the list. At the ORI, for example, director Chris Pascal says that a hands-on principal investigator (PI) who talks to junior scientists regularly and stresses the need to run experiments properly, rather than rushing out results, can make a big difference: "Get a PI like that and the risk of misconduct is much lower." Institutional polices that insist on good record-keeping are also essential, adds Pascal.

Both pieces of advice seem straightforward,

Could identifying fraud risks create a *Minority Report*-style police state?

20TH CENTURY FOX/DREAMWORKS/KOBAL COLLECTION

©2007 Nature Publishing Group

TEK IMAGE/SPL

Pressure cooker: the competition between and within biomedical research labs can be intense.

but neither are followed as much as they should be. A 2003 ORI survey[5] concluded, for example, that one in four lab heads did not take their supervisory roles seriously enough. Many institutions also fail to enforce data-management policies, says Pascal. That is worrying, as poor record-keeping was present in almost 40% of more than 550 misconduct cases studied in a survey published this month[6].

Would these actions address all the causes of misconduct? Almost all scientists have felt pressure at some point in their careers, yet the surveys suggest that the majority do not commit even minor misdeeds. "It doesn't push everyone over the edge," says Nicholas Steneck, a historian at the University of Michigan in Ann Arbor who has worked on misconduct policies for the ORI and other organizations. "It comes down to individuals."

Here the research is more controversial. Alongside environmental factors, the Oklahoma group has looked for links between personality traits and ethical decisions. For each of the four areas looked at, from data management to experimental practice, the team found that subjects with high ratings for narcissism returned low scores. A sense of entitlement — 'I'm owed this result because of my hard work' — also predisposes researchers to misconduct. But so does having a trusting view of others.

Such results might improve understanding, but the potential for abuse worries some. Martinson says the situation reminds him of the science-fiction movie *Minority Report*, in which premonitions are used to apprehend individuals thought to be future criminals. Universities might, for example, choose to reduce misconduct risk by screening for traits such as narcissism in potential employees. "What you can do with this is frightening," says Martinson. "It doesn't lead to positive social control. At the extreme it leads to a police state."

That need not be case, argues Stephen Murphy, a PhD student who has worked on several of the Oklahoma studies. Rather than weed out individuals, personality work can inform research-integrity courses. Asked about ethical decisions, for example, most researchers declare themselves more ethical than their colleagues. Illustrating this by giving scientists a questionnaire and then sharing the results with a group is a powerful way of showing researchers that they are more flawed than they think, says Murphy. Those with narcissistic tendencies would not be treated any differently, but they might benefit more from such an exercise.

PHOTOALTO/ALAMY

OPINION
Read the musings of *Nature*'s columnists online.
www.nature.com/news

Over the edge

Yet even when knowledge of individual and environmental factors has been plugged into integrity training, there remains a bigger question about the way science is run. Many scientists work under enormous pressure, even in an era of relatively generous funding. Career paths in science exacerbate the situation. Unlike many other professions, scientists must constantly prove themselves by publishing papers. Success also depends not just on the view of a few immediate colleagues, but on that of the whole field. "It's about building a reputation," says Martinson. "In science that's the coin of the realm."

The case of obesity expert Eric Poehlman shows how these factors can push people over the edge. During his trial last year, when he was sentenced to a year in jail after admitting falsifying data in papers and grant applications, he said: "The structure…created pressures which I should have, but was not able to, stand up to. I saw my job and my laboratory as expendable if I were not able to produce."

The similarities between Poehlman's testimony and that of many other fraudsters point to factors that institutions can tackle. The problem is that many of the risk factors for misconduct also seem to be what makes for good science. Most would agree that competition is needed to allocate over-subscribed research funds appropriately, as well as to push individuals to evaluate their ideas. So even with the best research environment and training, fraud is unlikely to disappear. "It's the dark side of competition," says Martinson. While the pressure remains, so will some level of misconduct. ■

1. Cyranoski, D. *Nature* **438,** 1056–1057 (2005).
2. Martinson, B. C., Anderson, M. S. & de Vries, R. *Nature* **435,** 737–738 (2005).
3. Mumford, M. D. *et al. Ethics Behav.* (submitted).
4. Martinson, B. C., Anderson, M. S., Crain, A. L. & de Vries, R. *J. Empirical Res. Hum. Res. Ethics* **1,** 51–66 (2006).
5. Rodbard, D. *et al. Survey of Research Integrity Measures Utilized in Biomedical Research Laboratories* (US DHHS, Rockville, Maryland, 2003); available at http://ori.dhhs.gov/documents/research/intergity_measures_final_report_11_07_03.pdf.
6. Wilson, K., Schreier, A., Griffin, A. & Resnik, D. *Account. Res.* **14,** 57–71 (2007).

©2007 Nature Publishing Group

Where are they now?

Nature catches up with some past fraud investigations — and finds that, whether researchers are found to be guilty or innocent, the wounds are slow to heal.

JON SUDBØ

NORWEGIAN RADIUM HOSP.

Jon Sudbø, a doctor and researcher at the Norwegian Radium Hospital in Oslo, confessed in January 2006 to fraud on an extraordinary scale (see *Nature* **439**, 248–249; 2006). He admitted creating some 900 fictitious patients in a 2005 *Lancet* paper on the effect of certain painkillers on oral cancer risk in smokers, since retracted. An independent commission of inquiry was set up to discern the details of the fraud; last June it reported that the bulk of Sudbø's 30-plus publications were invalid because of the fabrication and manipulation of data.

Sudbø resigned the day after the commission released its report and was stripped of his medical and discipline degrees late last year. He is reportedly still in Oslo with his family but not working. Richard Horton, editor of *The Lancet*, co-hosted a meeting in Norway last month to establish what could be learnt from the case. He thinks that because Sudbø used public funds to pay for fraudulent work, he will probably face criminal proceedings. **L.O.-S.**

THEREZA IMANISHI-KARI

A 1986 paper in *Cell* on immunity in transgenic mice acquired international fame — not because of what it reported, but because of the slew of investigations it spurred. Among the paper's authors were Thereza Imanishi-Kari and Nobel laureate David Baltimore, then both at the Massachusetts Institute of Technology. A postdoc in Imanishi-Kari's lab accused her of falsifying data in the *Cell* paper. The decade-long saga that followed involved a series of investigations, including one by the Energy and Commerce Congressional committee. Imanishi-Kari was cleared of all charges in 1996.

After doing no research for ten years, Imanishi-Kari reorganized a research programme at Tufts University in Medford, Massachusetts. Although pleased with the response of colleagues to her publications, she feels the case still overshadows her grant applications. "Being a woman in science is one kind of handicap; being wrongly accused is another," she says.

Henry Wortis, one of Imanishi-Kari's collaborators at Tufts, says people should judge her by her current work, published in leading journals such as *Immunity*. "That's the quality of work she's been unable to do for ten years," he says. "It's been a considerable loss to research." **L.O.-S.**

JAN HENDRIK SCHÖN

BELL LABS

The fraud that came to light in the spring of 2002 is by far the most serious to strike physics in recent years. Jan Hendrik Schön was a rising star at Lucent Technologies' prestigious Bell Laboratories in Murray Hill, New Jersey, and author of a string of papers in *Nature* and *Science*. His reputation collapsed following a Bell Labs report that found Schön guilty of falsifying data in at least 16 papers on the electronic properties of new materials, such as organic films that might one day replace silicon-based electronics.

Schön's whereabouts are now unknown, and he has declined to comment on his research while investigations are continuing (the University of Konstanz in Germany is still assessing the work he did to earn his PhD there in the mid-1990s). But shock waves from the case have hit physicists, mainly in the shape of new research guidelines. Some of Schön's co-authors were criticized for not spotting their colleague's misconduct, prompting the American Physical Society to issue new rules stating that co-authors should be accountable for important data in papers they sign off. **J.G.**

ANDERS PAPE MØLLER

T. A. MOUSSEAU

In late 2003, the Danish Committees on Scientific Dishonesty concluded that Anders Pape Møller, one of the world's leading experts in behavioural ecology, was guilty of 'scientific dishonesty'. A former colleague, Jørgen Rabøl, had accused Møller of using flawed data in a paper published in the journal *Oikos*. The paper has since been retracted.

Unconvinced, the National Council for Scientific Research in Paris, where Møller was and is still based, appointed an independent 'committee of wise men' to examine the case. In late 2004, the committee concluded that Møller was innocent: although it found that good lab practices hadn't been followed, it could not prove intent to commit scientific fraud. Møller did not respond to invitations for interview.

An impressive publication list in the past two years suggests that Møller's output has not been dimmed. But a collaborator, Timothy Mousseau at the University of South Carolina in Columbia, says that the consequences of the case rumble on. "Even though a complete investigation found no evidence, people get emotionally caught up in these kinds of interactions. There is no easy way to clear your name." **L.O.-S.**

©2007 Nature Publishing Group

BAE JAE-MAN/YONHAP/AP

WOO SUK HWANG
Follow the story of the infamous South Korean stem-cell researcher online.
www.nature.com/news/specials/hwang

WOO SUK HWANG

S.-H. YOU/REUTERS

South Korea's stem-cell researcher Woo Suk Hwang is arguably the highest-profile scientific fraudster ever. In two hugely cited papers in *Science* in 2004 and 2005, Hwang claimed to have derived stem cells from cloned human embryos. But on 10 January 2006, Seoul National University announced after a four-week investigation that all of his claims were fabricated.

The legal issues have been harder to clear up than the scientific ones. Prosecutors started an investigation the day that the university report was released, and delivered their own report on 12 May 2006. They charged Hwang with fraud, embezzlement, and violation of the country's bioethics law. But more than half-a-dozen hearings later, the court has still not reached a conclusion. The proceedings have not prevented Hwang from doing research: he is now carrying out animal cloning experiments at his new laboratory in Yongjin, 50 kilometres south of Seoul. **D.C.**

SHINICHI FUJIMURA

In the 1980s and 1990s, amateur archaeologist Shinichi Fujimura was known throughout Japan for his 'divine hands', thanks to his discovering a string of stone artefacts that dated back to 700,000 years ago. Then in 2000, Japanese newspaper *Mainichi Shimbun* published photos of Fujimura digging holes to bury tools that he later announced as major finds, and he confessed to fraud.

The Japanese Archaeological Association released the results of its final investigation in 2004. It concluded that all of the 168 sites dug by Fujimura had been faked — forcing Japanese history textbooks to rewrite their descriptions of the Palaeolithic period. The Tohoku Paleolithic Institute, where Fujimura held the post of deputy director, was also dissolved in 2004.

Fujimura, aged 56 years, was hospitalized for mental illness, and then reportedly remarried, changed his name and is now living quietly in a small town near the Pacific coast. **I.F.**

LUK VAN PARIJS

E. QUINN/CORBIS

Once considered an exciting prospect in immunology and RNA interference, Luk Van Parijs was sacked from the Massachusetts Institute of Technology (MIT) in Cambridge in October 2005 after admitting that he had altered data in a published paper, unpublished manuscripts and grant applications. The institute inquiry that prompted the confession was sparked by allegations from some of his lab members.

Surprisingly, Van Parijs has co-authored a publication since the scandal, on a paper published in *The Prostate* in February 2006. The study's first author — urologist Qiang Zhang of Northwestern University in Chicago, Illinois — says that he and his co-authors included Van Parijs as an author because he had provided helpful ideas when giving a seminar in Zhang's department a few years earlier. He and other collaborators contacted by *Nature* have not communicated with Van Parijs since he left MIT. **L.O.-S.**

BENGÜ SEZEN

In March 2006, Dalibor Sames, a chemist at Columbia University in New York, withdrew two papers and part of a third from the *Journal of the American Chemical Society* (*JACS*). Work done by graduate student Bengü Sezen, he said, couldn't be reproduced. The work was in the field of carbon–hydrogen bond functionalization, which aims to selectively break bonds within a molecule.

The case got a lot of coverage on the ever-lively chemistry blogs, and a few press outlets mentioned the issue, including *Nature* (**440,** 390–391; 2006).

Sames is not commenting further until an inquiry is completed, and Columbia University says that it is against its policy to "comment on the existence or non-existence of any internal investigation into allegations of research misconduct". But Sezen has vigorously defended herself in e-mails to the editor of *JACS* and the press. **E.M.**

Lucy Odling-Smee, Jim Giles, Ichiko Fuyuno, David Cyranoski, Emma Marris

How much fraud?

Statistics relating to the prevalence of scientific fraud and misconduct are sparse. But here are a few.

UNITED STATES

Of 3,247 researchers funded by the National Institutes of Health who responded to a survey, 1.4% admitted to plagiarism and 0.3% admitted to falsification of data. 15.5% of responders said that they had changed the design, methodology or results of a study in response to pressure from a funding source (B. C. Martinson *et al. Nature* **435,** 737–738; 2005).

In 2005, the Office of Research Integrity received 265 allegations of misconduct. Of 22 cases that it closed in that year, research misconduct was found in eight. Of these eight cases, three were for falsification; two involved falsification and fabrication; two involved fabrication; and one involved plagiarism.

UNITED KINGDOM

Of 13 cases of suspected fraud brought to the Committee on Publication Ethics (COPE) between 1998 and 2003, seven were given up because of lack of documentation or lack of a response from the author or institution.

JAPAN

Of 1,323 research institutes that responded to a Science Council of Japan survey in 2006, only 13.4% had guidelines on what behaviours are defined as misconduct and how to act if misconduct is reported. Just 12.4% of institutes had discussed dishonest behaviour with their researchers in the past ten years.

GERMANY

The ombudsmen of the German Research Foundation (DFG) were called in 45 cases of misconduct in 2004, and took up 36 of them. Of these 36 cases, misconduct was found in one investigation.

DENMARK

Of 25 investigations that were concluded by the Danish Committees on Scientific Dishonesty between 2003 and 2005, dishonest behaviour was found in two of them.

©2007 Nature Publishing Group

Plan to standardize federal assessments discarded

The US National Academy of Sciences has rebuffed a White House plan to change the way that the federal government assesses the risks posed by chemicals and other potential hazards.

On 11 January, a panel chaired by John Ahearne, head of the ethics programme at Sigma Xi in Research Triangle Park, North Carolina, ruled unanimously that the plan was "fundamentally flawed" and ought to be withdrawn. The plan included definitions of risk assessment and proposed standards for its uniform implementation by federal agencies.

Environmental groups welcomed the academy report: they had painted the plan — issued early last year by the White House Office of Management and Budget — as a back-door effort to loosen environmental regulation (see *Nature* 442, 242–243; 2006). White House officials said afterwards that the rules would not now be issued in their existing form.

India launches space recovery capsule

In what is seen as a prelude to a proposed manned space-flight programme (see *Nature* 444, 255; 2006), India has successfully launched its first recoverable space capsule.

The 550-kg capsule, which went up on 10 January, is expected to remain in orbit for about two weeks before parachuting back to Earth. It will help the Indian Space Research Organisation to evaluate re-entry and thermal protection technologies needed for manned flights, officials say. The craft also carries two experiments on making alloys in zero-gravity conditions.

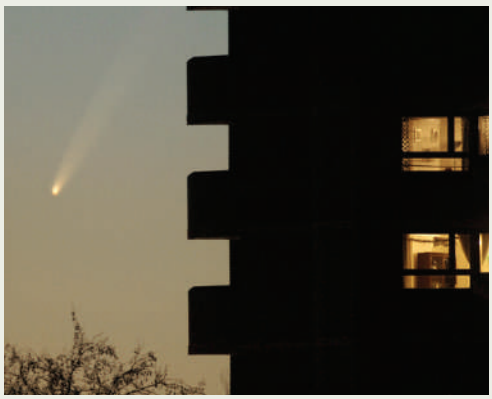
The capsule and a 680-kg high-resolution reconnaissance satellite were among the four payloads on board India's PSLV rocket, which took off from the Sriharikota launch pad on the country's east coast. A 6-kg nanosatellite from Argentina and a 56-kg microsatellite from Indonesia were also on board.

India's first recoverable space capsule is launched.

BABU/REUTERS

The night the comet came to town

Comet McNaught (right), named after Australian astronomer Robert McNaught (see www.nature.com/news), illuminated the streets of London on the evening of 10 January. Researchers at NASA's Solar Terrestrial Relations Observatory (STEREO) and the international Solar and Heliospheric Observatory (SOHO) scrambled to gain insight into its composition. "It's certainly the brightest comet seen from Earth in the past 30 years, perhaps even 40 years," says Jonathan Shanklin, director of the British Astronomical Association Comet Section.



A. MACDONALD/REX FEATURES

Gates grant to coordinate nations' public health

An international organization dedicated to building better links between national public-health agencies has received a US$20-million, five-year grant from the Bill & Melinda Gates Foundation.

The International Association of National Public Health Institutes aims to help poor nations combat health problems such as AIDS and bird flu by improving global communication between public-health officials. Since it was formed in January 2006 with a pilot grant, also from the Gates Foundation, it has attracted 39 members, including national public-health agencies from rich and poor nations.

The Gates grant was awarded jointly to Emory University in Atlanta, Georgia — home of Jeffrey Koplan, a former head of the US Centers for Disease Control and Prevention — and the Finnish National Public Health Institute. The association is the brainchild of Koplan, now its president, and his colleagues.

Whistleblower faces sacking by US government

A biologist employed by the US federal government to protect wildlife and the environment in Arizona is fighting for his career after his bosses sought to fire him for allegedly making unauthorized contact with outside groups.

Rex Wahl, at the US Bureau of Reclamation in Yuma, has been given 30 days from 9 January to respond to orders that he should be sacked for sending e-mails to other governmental agencies and to an environmental group. Bureau officials have described the communications as "conduct unbecoming a federal employee".

Wahl's lawyer, Paula Dinerstein of Public Employees for Environmental Responsibility in Washington DC, says that Wahl is being subjected to retaliation for properly upholding environmental laws, and should be entitled to legal protection as a whistleblower.

Korean stem-cell fraud claims another victim

More than a year has passed since the South Korean biologist Woo Suk Hwang was charged with fabricating data in highly cited papers about human cloning. Now, a former colleague has been found guilty of falsifying results and lying to officials investigating the case.

Jong Hyuk Park worked with Hwang in Seoul until 2004, when he joined the lab of Gerald Schatten, a stem-cell biologist at the University of Pittsburgh in Pennsylvania and a co-author on some of Hwang's key papers. Park's work came under scrutiny last January when Schatten alerted the university to concerns about a paper that Park planned to submit to *Nature*.

The university's investigation, released on 9 January, concludes that Park intentionally falsified two figures in the paper. He then "repeatedly misrepresented" the data to Pittsburgh officials investigating the case and deleted records from the lab's server. Park left the university shortly after the allegations surfaced and is believed to have returned to South Korea.

***NATURE* NEWS INTERNSHIP**

Nature is seeking an intern reporter to work full-time in either its Washington DC or London office starting July 2007 for six months. Applicants should be self-starting and have a keen news sense. This is a paid position.

The intern will write news and news features for *Nature* as well as online news for *Nature*'s website, news@nature.com.

Please e-mail a covering letter, your CV or resume, and three clips to both Alexandra Witze, senior news and features editor (a.witze@naturedc.com), and Nicola Jones, online news editor (n.jones@nature.com), by 9 February. Put 'internship application' in the subject line.

©2007 Nature Publishing Group

Climate of opportunity

With the shift of power in the US Congress comes an chance to re-engage in the debate over climate change. But the process will not be simple, says our new columnist **David Goldston**.

The 110th Congress convened in Washington DC on 4 January, with the Democrats in control of both chambers for the first time in a dozen years. The banging of the gavel by the new Speaker of the House, Nancy Pelosi, was in effect the starting gun for what is likely to be two years of political strife, as both Democrats and Republicans race towards the 2008 congressional and presidential elections.

The political and legislative manoeuvring is likely to be especially fierce because no one knows who holds the upper hand. It is true that technically the Democrats have the advantage. But, given their tenuous, two-seat majority in the Senate and the ideological diversity of their caucus in the House of Representatives, to what extent will they really be able to control the congressional agenda? Washington feels especially unsettled now, with a weakened president, an untested Congress, and a 2008 presidential primary process that is up for grabs in both parties.

But none of the manoeuvring is likely to yield significant changes for either the conduct or application of science in the near term.

To take one obvious example, the budget outlook for science is not likely to be appreciably different — the two parties support similar funding levels for research. (I'll have more to say on that next month, when President Bush releases his proposed budget for the next fiscal year.)

And the debate over embryonic stem-cell research is likely to be a rerun. The House has already passed a bill, identical to one passed last year in both chambers, easing the federal prohibition on research using new stem-cell lines. The Senate is likely to take up this or an amended version — but either way, the president is expected to veto it, just as he did its predecessor. The exercise will have more to do with making sure voters know where individual politicians stand on the issue than with changing stem-cell policy.

But Democratic control of Congress is likely to jump-start one genuine debate — that on policies to address climate change. Republican leaders had shut off all such debate in the House, and policy discussions have long been idling in the Senate, where the loudest voice was Senator James Inhofe (Republican, Oklahoma), with his show

PARTY OF ONE

trials to try to prove that global warming is a 'hoax.' Now, bipartisan efforts are under way to cobble together a regulatory bill that could make a dent in US greenhouse-gas emissions.

Such a task will not be easy. First, the key Senate players — Democrats Joe Lieberman of Connecticut, Barbara Boxer of California and Jeff Bingaman of New Mexico, and Republicans John McCain of Arizona and Pete Domenici of New Mexico — will have to figure out how to work together after years of offering up competing approaches.

Then they will have to decide if they want to compromise enough with the rest of their colleagues for the Senate to pass a bill, or whether they would be satisfied with building momentum for a more ambitious measure. The prospect of real action on climate change might actually make it harder to garner votes for a far-reaching bill. In past Congresses, some senators were willing to vote for a greenhouse-gas bill put forth by Senators McCain and Lieberman; such a vote signalled that they wanted to do something about climate change, but they didn't have to worry about the details because the bill clearly wasn't going anywhere with the Republicans in control of Congress.

Environmental groups are now in the process of making similar strategic calculations. Should they attack any measure that doesn't place a strict economy-wide cap on greenhouse-gas emissions, for fear that half-steps will sap political momentum? Or would such a move alienate their political allies as too absolutist, thereby slowing momentum?

Similarly, industry groups, which are all over the map on whether and how to regulate greenhouse-gas emissions, will have to decide where to draw their battle lines. They will no doubt face pressure not to give ground from Republican leaders, who, after all, still control the White House and may be back in power in Congress in two years.

And House Democrats have yet to embrace climate change as part of their agenda, fearing that it will fracture their fragile caucus, which features a greater number of politically vulnerable Democrats from oil- and coal-producing regions than the Senate.

But the biggest wild card of all on the climate issue — and the one that makes it so hard for any of the players to figure out how to play their hands — is the public. Polls generally show that the American public thinks climate change is real and wants action to counter it. But the depth of public understanding and support for action are far from clear.

Green politicians and groups have generally taken the line that environmental improvements can be made without personal sacrifice. Al Gore's movie *An Inconvenient Truth* seems to go out of its way to play down the difficulties in controlling climate change, to ensure that no one leaves the film thinking that taking action would be futile.

"The biggest wild card on the climate issue is the American public."

A serious debate in the Senate should increase the media and public focus on what is really at stake in regulating greenhouse gases. That might allow pollsters to get a firmer sense of where the American public stands on policy options — a prerequisite for any serious action on climate change. Given that several Congresses usually have to work on major legislation for years before it becomes law, engaging the public may be the most important result of any debate over the next two years. And the complexity of that policy discussion will make the previous congressional debate over whether climate change even exists seem like child's play. ■

David Goldston is a visiting lecturer and practitioner-in-residence at Princeton University's Woodrow Wilson School of Public and International Affairs. He served as chief of staff for the House Committee on Science between 2001 and 2006.

See Editorial, page 230

©2007 Nature Publishing Group

BUSINESS

P. SAKUMA/AP

Apple chief executive Steve Jobs has been accused of benefiting from backdated stock options.

Turning back the clock

Technology companies are under investigation for manipulating the timing of senior employees' stock options. **Heidi Ledford** reports.

Last week, on a sunny Tuesday morning in San Francisco, California, fans of Apple Macs queued up outside the convention centre to witness chief executive Steve Jobs' keynote address at the 2007 Macworld Expo. Jobs was due to speak at 9 in the morning; the line was already growing around midnight.

The annual Expo is a celebration of all things Apple. But attendees were not just queuing to see the new iPhone. They came for the pageantry and to witness Jobs' famous enthusiasm first hand.

Jobs has been in the spotlight for the wrong reasons this year, however — for his possible involvement in the manipulation of the timing of stock options at Apple.

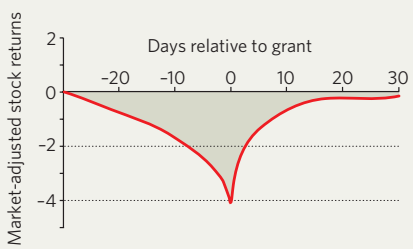
Stock options are awards of equity that senior employees can sell long after the date they were granted. But Apple, along with other US companies, now stands accused of fixing the dates of its stock-options grants retrospectively, so that their award date marks a low point in the share price. That means that when the option is exercised, the awardee can buy at this low price and sell at a tidy profit.

Apple, now the defendant in a lawsuit brought against it by shareholders, has already confessed to backdating stock options, and Jobs knew about the practice — he even recommended a few good dates to use. But whether or not Jobs benefited personally remains an open question.

Backdating is illegal only if the company does not account for it in its earnings reports. But it is now evident that many companies, including Apple, have not followed this procedure. In the past year alone, more than 120 similar investigations have been launched, and lawsuits have been brought against 22 companies for securities fraud.

For some companies, the consequences are minor. Affymetrix, a gene-chip company in Santa Clara, California, launched an internal investigation in response to a shareholder lawsuit. It eventually concluded that despite a few isolated "documentation lapses", there was no pattern of stock-option backdating. The lawsuit is still pending, but Quintin Lai, an analyst at the Milwaukee-based financial-advice firm Robert W. Baird, doubts that the episode will have a major impact on the company.

RETURNS AROUND STOCK-OPTION GRANTS (1992–2002)



SOURCE: E. LIE

But the consequences can sometimes be severe. "In the worst-case situation," says Joseph Grundfest, a securities law professor at Stanford University in California, "you lose your senior management team, the company spends tens of millions of dollars on an investigation, and the stock gets trashed." Executives at Comverse, a telecommunications firm based in New York, for example, have been charged with criminal fraud. One of the executives agreed to pay US$3 million to settle the case with the Securities and Exchange Commission (SEC) — the body that regulates public companies in the United States.

The allegations started to gather steam after Erik Lie, an economist at the University of Iowa in Iowa City, sent a paper he wrote on the topic (*Management Sci.* **51,** 802–812; 2005) to the SEC. The paper described a peculiar trend — a graph of average stock prices before and after a stock-option grant was shaped like a 'v', with the trough on the day of the award (see graph). "The results suggest that at least some of the awards are timed retroactively," Lie wrote.

The SEC's investigations became public in spring 2006. So far, more than half the investigations are in sectors such as computer software and biotechnology, in which profits are often low and senior executives look to stock options for their reward. "In the high-tech sector, the currency of the realm was stock options," says James Cox, a law professor at Duke Law School in Durham, North Carolina. "It was a very seductive environment for backdating to prevail."

Most investigations revolve around awards granted before 2002, when the SEC tightened disclosure requirements in response to the Enron scandal, and required that options be reported within two days of their starting date. That gave firms a much smaller window from which to select a stock price, and Lie's data show that the prevalence of backdating of stock options retreated sharply after that time.

But Lie says that the practice was still alive in 2005, with 13% of stock-option grants still filed late, and the SEC does not have the staff to pursue late compliers.

And as painful as it may prove to be for Jobs and others who have been accused, the Apple scandal is unlikely to rock corporate America as the Enron one did five years ago. Grundfest says that enhanced disclosure requirements have already drawn the sting from the issue. "The word is out," as Grundfest puts it: "Thou shalt not backdate." Companies have now found other ways to reward key employees, such as restricted stock. Only time will tell whether these methods are also being manipulated to gain an advantage over ordinary shareholders. ■

©2007 Nature Publishing Group

WORKING OUT THE BUGS

Programming a robot to think like an insect is tough, finds **Alison Abbott**, but it could help breed machines as manoeuvrable as flies.

Tarry II looks like a robot, sounds like a robot, and walks like an insect. He sprouts a tangle of wires, and the mechanical joints on his six legs emit a metallic creak with every step. But he strides determinedly across the lab with the steady gait of a fly marching towards rotting fruit.

The strutting machine is the work of Roland Strauss at the University of Würzburg and robotics colleagues elsewhere in Germany. Strauss acquired a zeal for mechanics as a schoolboy, when he won a national prize for designing electronics that could imitate a crab's visual system. But his real passion was always biology, and his goal today is to understand how behaviour involving movement is controlled in insects. He has recruited Tarry II and a small army of other biorobots to help.

Although our encounters with flies often leave an impression of aimless and irritating meandering, these tiny creatures' decisions are just as purposeful as those of other animals. A fly scans its environment with eyes and antennae, processes this information in its brain and then makes a decision, perhaps to turn away from potential danger or hurry towards food. Strauss aims to tease apart the complex of brain circuits that coordinates such movements. He hopes to identify broad principles about how the brain directs these behaviours, which might also apply to other animals including, perhaps, ourselves.

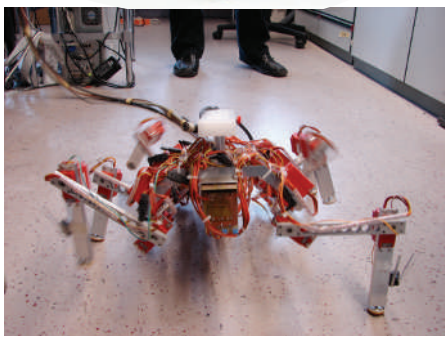
To help, Strauss and a handful of other insect biologists have turned to robotics experts. By programming simple robots to react to stimuli and move in particular ways, they can test biological hypotheses about which neural networks an insect uses to navigate.

And the biologists hope to return the favour. The algorithms they use to direct their biorobots could in the future help design smarter and more agile robots, capable of overcoming many barriers without direction from humans. (Videos of both insect-like robots and insects whose movement behaviour has been experimentally manipulated can be seen in this feature on the *Nature* website.)

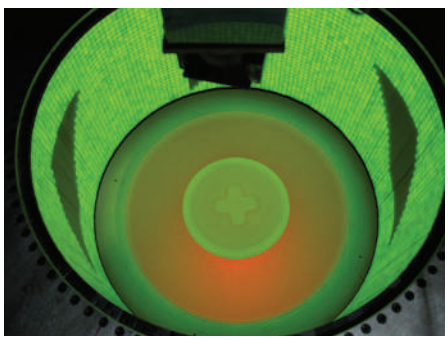
If only the Mars rovers had been more like cockroaches, sigh insect biologists, they might have been able to extricate themselves from the sand dunes and rocks on which they have occasionally come a cropper and had to be carefully steered to safety by their human controllers. "We are very happy if what we learn from nature can be put to use to make better robots," Strauss says.

A large amount of insect movement occurs without guidance from the brain. It is instead directed by circuits of nerves in the nerve cord, which extends from the insect's head into its thorax and is equivalent to the vertebrate spinal cord. If you decapitate a fly it will stop moving because the brain no longer provides the 'go' signal. But if you drop a neurotransmitter such as octopamine directly onto its thoracic nerve cord, then it will start to walk around like — well, like a headless chicken[1]. The fly can even be induced to carry out more complex and ghoulish behaviours, such as grooming eyes that are no longer there. These basic movement programmes are well studied and have been transferred to robots. Tarry II's prototype, designed by biorobotics pioneer Holk Cruse at the University of Bielefeld, has been walking with the confident coordination of a decapitated stick insect for more than a decade (see video 1 on *Nature*'s website).

R. STRAUSS



A. ABBOTT



R. STRAUSS

Insect-mimics Dro-o-boT (top) and Tarry II (middle), and a chamber for studying fly movement (bottom).

Executive decision

But for the cleverer stuff — deciding when to move, at what speed and in which direction — the insect recruits its brain. Just as we do. And just as robots should, but cannot, because current algorithms do not to provide them with this level of sophisticated and autonomous decision making.

Like other animals, an insect's brain contains specialized circuits responsible for controlling particular aspects of behaviour, spanning defined anatomical areas. So biologists have begun to decipher which particular region is involved in directing each type of movement by destroying that area and watching what happens. Strauss likes to work with the fruitfly *Drosophila melanogaster* because genetic tools are available to selectively damage very specific brain areas. Researchers can, for example, introduce random mutations into the fly's genome and then screen for flies that walk or move abnormally. In many of these, a mutation will have interfered with the structure of a particular brain region.

Using such methods, Strauss has accumulated a collection of several hundred mutants with aberrant movement behaviours. A large proportion have damage to a brain structure called the central complex, a collection of components sitting between the two hemispheres of an insect's brain known to be critical for

©2007 Nature Publishing Group

DK LIMITED/CORBIS

executive control of movement. Mutants with a damaged central complex might walk more slowly, or spiral their way towards an object of desire instead of approaching it in a straight line. Over the past ten years, Strauss has dissected out exactly how different regions of this complex cause specific movement or orientation problems, and he is now starting to see if he can translate these functions successfully into algorithms for Tarry II as well as for a newer, wheeled machine called Dro-o-boT.

Already these robots have supported his hypothesis that flies use parallax motion to gauge distances — the motion familiar to train passengers, who see close objects rushing past while distant ones seem to move slowly. The ability to gauge distance is critical for flies; they approach stationary objects in the hope that they may turn out to be safe havens or sources of food, and to save energy they always approach the closest first.

To monitor how flies move when they see objects, Strauss built a virtual-reality environment which looks like a a nightclub for flies. The setup consists of a small cylinder with a glass floor, and walls packed with light-emitting diodes that can be programmed to create geometric patterns of light and dark. The movements of a fly placed in the centre of the floor are videoed from above, and each footfall is recorded automatically from below. Test flies have their wings clipped so they cannot fly, and a tiny moat keeps them in the central arena because they are repelled by water.

Parallax view

Flies, it turns out, will move towards dark shapes created by unlit diodes as if they were real objects. To test his parallax-motion hypothesis, Strauss programmed his virtual-reality environment to automatically move the shapes according to the flies' movements. The flies did, as predicted, use parallax motion to identify the nearest 'object', which they approached[2]. But Strauss next wanted to test whether insects use parallax motion alone to judge distance or whether they also recognize what the objects actually are. For this, he turned to Dro-o-boT.

The researchers placed Dro-o-boT in a scaled-up, two-metre-square model of the virtual-reality environment. Dro-o-boT has an integrated 360° camera for eyes and was programmed to integrate visual information on the size of 'objects' with its own movements, so that it could recognize and wheel itself towards the dark spot that appeared closest (see video 2). Movement traces of Dro-o-boT confronted with four landmarks were virtually identical to those of a fly confronted with an equivalent four landmarks, showing that parallax motion is indeed sufficient to direct this type of movement.

In the past couple of years, Strauss's group has similarly programmed Tarry II so that it can imitate flies even more closely than the wheeled Dro-o-boT, by using parallax motion to identify and approach the closest object on its six legs. The robot can even follow the movements of its owner relative to marks on the more distant wall and change direction to move towards him (see video 3).

Having explored how a fly reacts to an object, Strauss is now starting to examine other movement behaviours, including walking speed. Such is the simplicity of a fly's brain that if two black squares appear directly opposite each other on the cylinder walls, it will walk towards one, then turn at the moat and pace back towards the other, lacking the intellectual wherewithal to choose a destination. It thus continues a seemingly pointless march until fatigue sets in (see video 4). This is known as the Buridan paradigm, after 'Buridan's ass', a thought experiment in which a donkey placed exactly between two identical piles of hay is destined to starve to death because it cannot decide which pile to choose. In another variation, the virtual-reality environment is set up with vertical lines and rotated, making the insect circle at the same speed as it continually tries to walk towards the nearest object.

"Robots are going to help us work out increasingly complex movement behaviours in insects."
— Robert Full

The fly's predictable behaviour in these paradigms is useful for experimentalists because it allows them to find mutants with brain damage that interferes with their turning, walking speed and resoluteness (the time before

©2007 Nature Publishing Group

they simply give up the march). Many of the mutants that performed badly in these two experiments turned out to have damage to a structure within the central complex called the protocerebral bridge, a string of nerves which Strauss has shown changes the swing speed of the leg and hence controls step length. In one of his mutants, appropriately called 'no bridge', the bridge is cut right through. Another, called 'tay bridge', after a Scottish bridge that collapsed disastrously in 1879, has a decided kink.

Normal flies take longer strides when their stride rate increases, to gather speed efficiently. But in Buridan's paradigm, tay bridge mutants did not increase stride length, walked at half the speed of normal flies, and gave up earlier. In the rotation paradigm, neither the no bridge nor the tay bridge mutants could keep up with the moving stripes.

Why robots need legs

Just as horses use different leg sequences when they trot, canter or gallop, six-legged insects have different gaits depending on their speed. On a smooth horizontal surface, a fruitfly normally adopts an alternating tripod gait, where the front and back leg on one side, and the middle leg on the opposite side, swing together in unison supported and balanced by the other three legs in a sturdy tripod. To move faster, it will increase both the frequency of steps and the speed of the swing to lengthen the stride. To turn, it reduces the length of the stride on the inside of the turning curve.

When a six-legged insect walks slowly, it adopts a tetrapod gait, where legs swing together in a type of wave from front to back and at least four legs are on the ground at the same time. Both gaits can stably accommodate objects underfoot — the insect simply adjusts its body position for balance.

Evolution has found different ways to achieve balance and locomotion in insects and vertebrates — and now robot designers must solve the same problem. Legged robots, with their stability and flexibility over uneven ground, offer clear advantages and some machines already copy their style. By using the tripod gait, for example, the Finnish Forest Walking Machine is unfazed by rough terrain that might defeat vehicles with caterpillar tracks. And the US Department of Defense report following the terrorist attacks on the World Trade Center commented that legged robots might have been better at locating bodies in the rubble. **A.A.**

Mind the gap

Strauss used Tarry II to understand in more detail why these mutant flies are so sluggish. Graduate student Simon Pick, now at the University of Ulm, Germany, programmed Tarry II to mimic the mutants so that the robot no longer increased stride length with stepping frequency. In doing so, he uncovered what robotics experts call an emergent property. He found the tay bridge way of walking was not just slow, but also inefficient — it consumed 8% more energy for a given speed than normal walking. "We didn't expect it, but the robot simulation made us realize that energy efficiency is an important evolutionary advantage for walking insects," says Strauss. The evolutionarily ancient brain of the stick insect, he notes, is not so sophisticated: the insect can vary only step frequency, not step size.

Strauss is now turning his attention to more complex movements: his current preoccupation is how flies climb across gaps. By filming flies as they confront gaps of varying widths, he has described distinct units of behaviour as the flies try first of all to assess the width of the gap, then initiate and complete the crossing or, alternatively, topple into the abyss.

When a fly stumbles upon a chasm, for example, it typically raises its front legs high above its head and performs far-reaching climbing movements as if it were feeling for a way across. If it fails to find a way over, a normal fly tries again and will eventually throw its body across to grasp the far edge. But Strauss has identified several mutant lines which consistently foul-up gap-crossing, and seem unable to improve. One fails to reach the other side correctly because it initiates leg-over-head movements too early. Another fails to even start the crossing, choosing instead to turn tail and walk away[3] (see video 5).

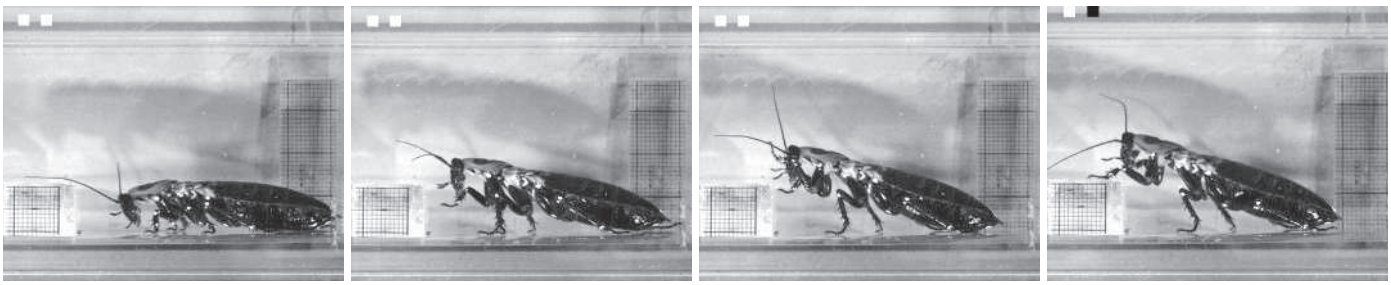
Strauss wants to use his robots to test how flies control their reaching when they negotiate gaps — but these complicated behaviours are difficult to model in machines. He is planning to fit Tarry II with attachments that would allow him to climb in the same way as flies need to when they cross a gap. "One might think about Velcro attachments and an environment with carpeted walls, or claw attachments with Styrofoam walls," he says.

Mutant fruitflies are a treasure trove for movement-behaviour researchers, but their minuscule brains do impose some experimental limitations. The cockroach brain, on the other hand, is 50 times larger, so researchers can inflict brain damage using surgery. Other advantages are a matter of taste. "Cockroaches can actually be quite pretty, I think," says neuroscientist Roy Ritzmann from Case Western Reserve University in Cleveland, Ohio, "and my sort [*Blaberus discoidalis*] don't really smell so bad."

About turn

In one series of experiments, Ritzmann removed the cockroaches' wings and severed the circumoesophageal connective, the part of the nervous system that separates the lower and upper parts of the brain. Following this procedure, cockroaches will walk until they drop, but cannot negotiate obstructions. When climbing an incline of 40–45°, for example, their legs might slip and they sometimes fall backwards in slapstick fashion (see video 6). This is because some of the brain-damaged insects are unable to control the height of their bodies and tend to raise their centre of mass too high, so adhesive pads on their legs no longer stick[4].

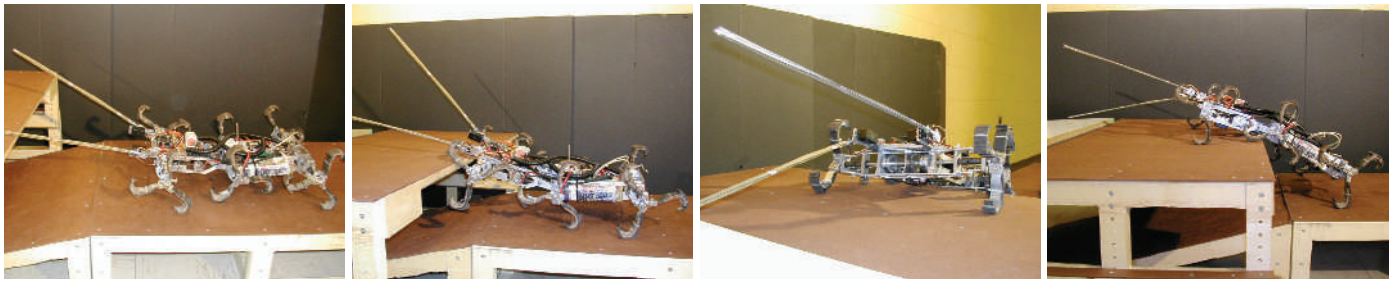
Like Strauss, Ritzmann wants to unravel the neural mechanisms his creepy-crawlies use when confronted with obstacles, behav-



R.E. RITZMANN

Up and over: a cockroach will climb over a shelf if its antennae touch the top of it first, behaviour that robots can mimic (see opposite page).

©2007 Nature Publishing Group



R. QUINN

The roach approach: a Whegs robot, moving on a cross between wheels and legs, clambers over a shelf using rules copied from cockroaches.

iour which is coordinated by the central body complex. His students carefully exposed the brains of cockroaches and used tin foil to make cuts through the complex. They then identified ones that no longer turned correctly by videoing them in a transparent chamber and later dissected out the brains to identify exactly which part had been damaged.

Intact cockroaches turn to the left when their right antenna touches the wall, and vice versa. But few of Ritzmann's lesioned cockroaches could turn properly. Depending on where the cut was made, some animals could turn only in one direction, some circled continuously and others were disinclined to turn at all, crashing into the walls[5] (see video 7).

These disabled roaches have helped Ritzmann develop a hypothesis that the insect brain goes through three distinct mental steps when it confronts an obstacle: it must detect the block, decide that it would be appropriate to turn, and then decide which way to turn. Ritzmann now wants to test this hypothesis by programming these three steps into a robot; he is collaborating with colleague Roger Quinn in the university's biorobotics laboratory. Quinn's six-legged Robot III is modelled on the *Blaberus* cockroach and it already possesses the astonishing stability of a headless insect — give it a strong shove, and it will find its feet again almost instantly (see video 8). But this work is just beginning and as yet Robot III can't walk very well, let alone negotiate turns.

Using a particularly manoeuvrable breed of robots called Whegs, Quinn has tested another of Ritzmann's hypotheses about a behaviour at which cockroaches seem adroit — climbing over or slipping under a shelf.

On the shelf

Whegs, so called because they are propelled by a cross between wheels and legs, have the tripod gait of all six-legged insects (see 'Why robots need legs') and can flex their bodies thanks to a special body joint. Quinn used his robots to test the idea that a cockroach's decision to clamber over or crawl under the shelf depends on whether its antenna first brushes the shelf's top or its underside. This is hard to prove definitively in the insects because many types of information or neural circuits might be involved in the decision. But a robot, which can be programmed with a set of algorithms so that it reacts only to information from the antenna, can reveal whether this sensory input is sufficient to direct its behaviour. Quinn's student William Lewinger added suitably programmed antennae to the Whegs, allowing the robot to detect a shelf's top or bottom and trigger it to climb or crouch respectively. As predicted, the robot behaved just like the real roach (see video 9).

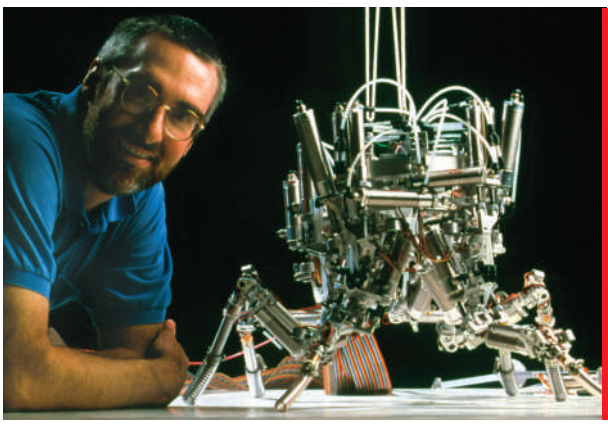
Insect biologists are eager to model ever more intricate types of insect behaviour in their robots, such as walking uphill or climbing, and some robots with this hardware capability already skulk in other labs. One six-legged machine called RiSE, built by a team at Carnegie Mellon University, Pittsburgh, can even climb trees. "Robots like these are going to help us work out increasingly complex movement behaviours in insects," says Robert Full at the University of California, Berkeley, whose studies on the mechanics of insects and other arthropods have inspired RiSE and other machines. But until these robots can be programmed with more sophisticated and autonomous software — precisely the directions that biologists are extracting from insect's brains — they cannot pass for true robotic insects (see video 10).

Insect biologists aren't the only ones anxious to tap into flies' intellect. Just a few of an insect's effortless navigational skills would be a boon for many of today's applied robots, which can negotiate obstacles only via human intervention and remote control. For this reason, space agencies such as NASA and the European Space Agency are watching developments in these insect labs with keen interest — and many robotics groups already include neurobiologists on their teams. "We think we may find it easier to go to higher control levels with neurobiological input," says Dirk Spenneberg of the University of Bremen, Germany, who works on autonomous robots including six- and eight-legged machines.

Quinn, for example, says he is working with NASA to determine whether Whegs could work well on the lunar surface, and has other collaborations exploring their use for surveillance, and search and rescue. And nothing would make insect biologists happier than seeing future generations of Tarry lookalikes confidently striding the canyons of Mars. ■

Alison Abbott is *Nature*'s senior European correspondent.

P. MENZEL/SPL

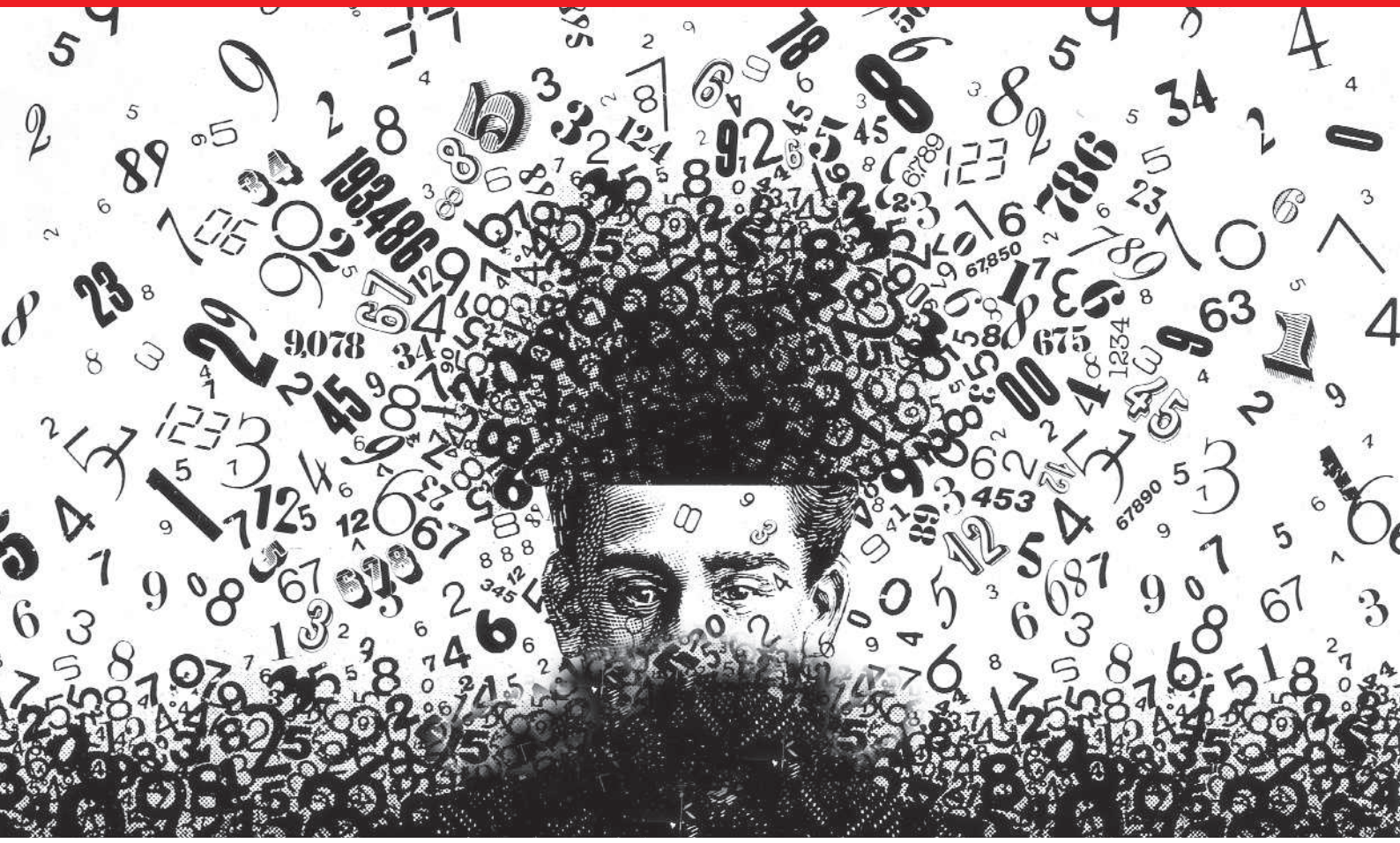


"Roger Quinn says he is working with NASA to design robots that can work on the lunar surface, and has other collaborations exploring their use for surveillance, and search and rescue."

1. Yellman, C., Tao, H., He, B. & Hirsh, J. *Proc. Natl Acad. Sci. USA* **94,** 4131–4136 (1997).
2. Schuster, S., Strauss, R. & Götz, K. G. *Curr. Biol.* **12,** 1591–1594 (2002).
3. Pick, S. & Strauss, R. *Curr. Biol.* **15,** 1473–1478 (2005).
4. Ritzmann, R. E., Pollack, A. J., Archinal, J., Ridgel, A. L. & Quinn, R. D. *J. Comp. Physiol. A* **191,** 253–264 (2005).
5. Ridgel, A., Alexander, B. E. & Ritzmann, R. E. *J. Comp. Physiol. A* doi:10.1007/s00359-006-0193-7 (2006).

©2007 Nature Publishing Group

NATURE|Vol 445|18 January 2007



IMAGES.COM/CORBIS

CONVICTION BY NUMBERS

Statistics have the power to trip everyone up — including judges and juries. Even when extra care is taken to get the numbers right in court, confusion often reigns. **Mark Buchanan** reports.

In March 2003, when nurse Lucia de Berk faced trial in a Dutch court for charges of murder and attempted murder, the statistical evidence against her seemed compelling. Investigators had identified a number of 'suspicious' deaths and near deaths in hospital wards in which de Berk had worked from 1999 to 2001, and records showed that she had been present when many of those events took place. The statistical expert testifying in the case, Henk Elffers, reported that the chance that her presence was mere coincidence was only 1 in 342 million.

On the basis of this number and on limited forensic evidence — traces of toxic substances found in two of the exhumed bodies — the court found de Berk guilty, and sentenced her to life in prison. But some Dutch mathematicians now say that the figure cited was incorrect, and that the case is a classic example of how statistical reasoning can go horribly wrong. "The magical power of the big number led everyone at an early stage to be totally convinced of Lucia's guilt," says mathematician Richard Gill of the University of Leiden in the Netherlands. "Then they went to work to confirm their theories."

A court of appeal later upheld de Berk's conviction, but an advisory judicial committee has now been appointed by the central office of public prosecutors to reassess the case. The committee's decision is expected later this year, and it could recommend that the case be reopened. But whatever the result, the case illustrates the ongoing difficulty of ensuring that courts use statistical reasoning properly. "This is a serious problem," says statistics professor Philip Dawid at University College London, "but there is no easy solution."

Learning from experience

The case also raises questions over whether courts have learned anything from past misuse of statistics. In a high-profile British case in 1999, Sally Clark was convicted of murdering her two small children based, at least in part, on faulty statistical reasoning by an expert witness — the paediatrician Roy Meadow. Dawid was invited to submit written evidence on the statistical arguments during Clark's first appeal. He was not, however, allowed to present oral testimony. "The lawyers and the judge argued that it was not rocket science, so a statistical expert was not needed."

The appeals court confirmed Clark's original conviction, but Dawid says that the written judgement showed that the jury had failed to grasp the relevant statistical issues. A second appeal in 2003 freed Clark, concluding that the jury might have been misled in part by the statistics presented in the original trial. Without a change in legal attitudes and procedures, mathematicians worry that statistical arguments will continue to be misinterpreted by the courts.

In de Berk's case, Gill and another Dutch mathematician, Peter Grünwald of the National Research Institute for Mathematics and Computer Science in Amsterdam, have submitted letters to the judicial committee arguing that the original figure of 1 in 342 million was incorrect — or, at best, irrelevant to the proceedings. De Berk first became a suspect when working at the Juliana Children's Hospital in The Hague. But because observations from this ward were the source of the initial suspicion, Gill and Grünwald say, those observations should not have been used in a calculation that tested the validity of the suspicion.

©2007 Nature Publishing Group

Elffers, of the Netherlands Institute for the Study of Crime and Law Enforcement in Leiden, combined the Juliana data with data from another hospital in which de Berk had previously worked to get his figure. Gill and Grünwald insist that the analysis was misleading. "It makes little sense to do formal hypothesis testing when the data themselves have suggested the hypothesis," says Gill. "The only safe thing is to go out and collect new independent data."

Gill's own calculation estimates that the probability that the correlation arose by chance is not 1 in 342 million, but a much smaller 1 in 48, or even as low as 1 in 5 — figures that are unlikely to meet the 'beyond reasonable doubt' needed for a criminal conviction. But Elffers defends his original calculations, arguing that he applied a factor that corrected for his using some of the data twice. "Everyone is aware that I applied a correction," he says.

Fact or fallacy?

Aside from this debate, equally important is how the court interpreted the number. Philosopher of science Ton Derksen of the University of Nijmegen, who has written a book that criticizes de Berk's conviction, argues that the court made an elementary statistical error known as the prosecutor's fallacy.

The court needs to weigh up two different explanations: murder or coincidence. The argument that the deaths were unlikely to have occurred by chance (whether 1 in 48 or 1 in 342 million) is not that meaningful on its own — for instance, the probability that ten murders would occur in the same hospital might be even more unlikely. What matters is the relative likelihood of the two explanations. However, the court was given an estimate for only the first scenario. Without additional information, says Derksen, Elffer's number is meaningless — and could easily be misinterpreted as a very small chance that de Berk is innocent.

To raise further doubt, other important statistics were neglected by the court. When de Berk worked at Juliana between 1999 and 2001, there were six unexplained deaths in her unit. The same unit, in a similar period before de Berk started working there, had seven unexplained deaths. "It seems very strange," says Grünwald, "that fewer people die when there is a serial killer around." Derksen says that the statistics comparing deaths before and after de Berk started work at the hospital were mentioned by her defence lawyers, but were not sufficiently emphasized to have any influence on the court.

Due process

This neglect illustrates a difference between legal and scientific processes. Although science aims to bring together all relevant evidence, this is not necessarily true with the law. David Kaye, an expert in statistics and the law at Arizona State University in Tempe, notes that lawyers have an incentive, and even a duty, to select the evidence that makes their case stronger. "What the judge ends up hearing often comes from the two extreme ends of the distribution," he says.

> "The magical power of the big number led everyone at an early stage to be totally convinced of de Berk's guilt."
> — Richard Gill

Procedures to correct such distortions are also lacking, even after a trial has reached a verdict. In the United States, written statistical arguments are often protected by court orders, and so are not available for review or correction. "The data pertaining to an individual deserve some protection," says statistical expert Joseph Gastwirth of George Washington University in Washington DC, but a summary of the expert reports should be made publicly available, he suggests.

Independent scientific comment of this kind occurred during the Clark case, but to unknown effect. In the 1999 Clark trial, Meadow testified that the chance of two infants from the same mother dying of Sudden Infant Death Syndrome (SIDS) was only 1 in 73 million. Two years later, after the first appeal, the Royal Statistical Society in London, condemned both this figure and its interpretation. The figure would be valid only if SIDS cases arise independently within families, the statement said, whereas there may be unknown genetic or environmental factors that predispose families to SIDS.

After the Clark case, the society established a working group, chaired by statistician Colin Aitken of the University of Edinburgh, to examine how courtroom use of statistics might be improved. In the immediate future, the society hopes to help provide continuing education to practising lawyers, judges and other legal practitioners, so that they can at least recognize the potential hazards in statistical reasoning. Ultimately, Aitken suggests, including statistics in the core curricula of law degrees will be more effective. "Things will not improve overnight," says Aitken, "but we are in it for the long haul."

A matter of opinion

In the United States, the Federal Judicial Center, an organization created by Congress to improve federal courts, has published a reference manual on the use of scientific evidence, with one chapter devoted to statistics. "But education is only palliative," says Kaye, who helped to write the statistics chapter. "I don't think there is any single way to ensure that statistics and other scientific evidence gets used accurately."

Although courts expect one simple answer, statisticians know that the result depends on how questions are framed and on assumptions tucked into the analysis, giving tremendous room for legal argument. Kaye recalls giving lawyers in one case what he thought was a crystal-clear explanation of a statistical argument. "Their response was 'Let's just put him on the stand, he will confuse everyone.'"

Indeed, the biggest practical challenge, some argue, lies in the unusually subtle nature of statistical reasoning, which research shows confuses experienced professionals, such as physicians, just as easily as the general public. In de Berk's case, even Elffers now suggests that with so much uncertainty swirling around the statistics, they should play no further role in the considerations. "The review committee should concentrate on non-statistical arguments," he says.

Those arguments are also a matter of dispute. Until the committee decides, de Berk languishes in prison, quite possibly because of mathematical errors. "I am so convinced that the statistics are wrong," says Gill, "that I am more inclined to believe in the incompetence of the entire process than in the existence of a serial killer." ■

Mark Buchanan is a science writer based in Europe.

K. WIGGLESWORTH/EMPICS

Sally Clark was released on appeal after questions were raised about statistical arguments related to her conviction for murder.

DE BERK'S FAMILY/DE TELEGRAAF

Lucia de Berk is still waiting to hear whether a judicial review of her conviction for murder will recommend that her case be reopened.

©2007 Nature Publishing Group

How a database of nuclear databases could help the effort to combat trafficking

SIR — We welcome the views expressed in the Commentary by M. May, J. Davis and R. Jeanloz on source attribution of nuclear materials ("Preparing for the worst", *Nature* **443,** 907–908; 2006), in particular the suggestion to establish an international databank of such material to support the interpretation of nuclear forensics analyses.

We would like to comment on a few points, in particular those where reference was made to the Institute for Transuranium Elements (ITU) in Karlsruhe, part of the European Commission's Joint Research Centre.

The ITU has been involved since the early 1990s in the analysis of seized nuclear materials from cases of illicit trafficking. Its database, mentioned in the Commentary, contains materials from several western European and Russian fuel manufacturers, and was established in collaboration with the Bochvar Institute in Moscow. For reasons of commercial sensitivity, its content is largely confidential. Nevertheless, we assess information requests on a case-by-case basis.

Most nuclear or radioactive material has been seized in quantities too small for the production of a nuclear device. However, it is conceivable that small quantities could be used by terrorists for a radiological dispersal device, or 'dirty bomb', which would also require source attribution.

As your Commentary mentions, the Nuclear Smuggling International Technical Working Group discusses scientific and technical aspects of nuclear forensics analyses. Its 2006 meeting discussed various options for information sharing, including the creation of a 'super' database containing information about other organizations' databases. This decentralized approach may be easier to implement than a single centralized database. Non-nuclear states also need to be involved in this effort.

We agree that incidents involving nuclear material call for a comprehensive response, which relies on an organizational and technical infrastructure. This includes 'crime scene' management to preserve forensic evidence — nuclear as well as traditional — and maintain the chain of custody.

Credible nuclear forensics is based on representative samples and on high-quality analyses carried out in specialized multidisciplinary laboratories. Some parameters (such as age, intended use and production mode) are self-explanatory. But others (such as impurities and pellet dimensions) can best be understood by comparison to references, which calls for information on a wide variety of materials or the ability to pose queries in databases. Source attribution is essential in the fight against illicit trafficking and in the prevention of malicious acts.

Klaus Luetzenkirchen, Klaus Mayer
Institute for Transuranium Elements,
Joint Research Centre, European Commission,
PO Box 2340 , Karlsruhe 76125, Germany

Time to give due weight to the 'carbon footprint' issue

SIR — The media are increasingly using the term 'carbon footprint' in articles about the need to mitigate climate change by reducing our carbon dioxide emissions. Footprints are spatial indicators, measured in hectares or square metres. The property that is often referred to as a carbon footprint is actually a 'carbon weight' of kilograms or tonnes per person or activity.

To improve public understanding of the issues surrounding climate change, it is necessary to be precise. Other 'footprints', such as the ecological or environmental footprint, convert resource consumption and waste production into spatial units. The term 'ecological footprint' was coined by William E. Rees, a planner at the University of British Columbia — who had previously used the term 'appropriated carrying capacity' — after a computer delivery man told him that the new machine, which took up less space than his old model, had a 'smaller footprint'.

As well as the media, many government agencies and environmental groups now use the expression 'carbon footprint'. Those who favour precision in such matters should perhaps campaign for it to be called 'carbon weight', or some similar term. That would avoid lasting confusion. Losing weight might even take on a whole new meaning.

Geoffrey Hammond
International Centre for the Environment,
and Department of Mechanical Engineering,
University of Bath, Bath BA2 7AY, UK

Moon and Venus as worthy of exploration as Mars

SIR — While joining in your celebration of achievements by Mars Global Surveyor and remaining Mars spacecraft (*Nature* **444,** 519 and 526–527; 2006), I disagree on priorities. The Editorial commends "the idea of a continuous presence" on Mars, with "various spacecraft calibrating and complementing each other's results". Yet the related News story seems to accept the view that "other missions might be easier to finish, such as Venus Express — there's a limit to how much useful data can be gathered by continuing to orbit Venus with the same instruments".

According to this 'science per dollar' principle, if continuous and intensive exploration is appropriate for Mars, then it is also appropriate for Venus, which is no more costly to reach and features a dynamically varying atmosphere and surface that have evolved in an intriguingly different way from Earth. The same principle also applies to the Moon. Although not dynamically changing, it holds valuable records of the early Earth's environment, and is reachable in a fraction of the time it takes to reach the planets.

Let us hope that the new US Congress and its international counterparts take the best ideas from all sources and assemble a balanced programme of lunar and planetary exploration for the coming decade.

Curt Covey
1010 Lynn Street, Livermore,
California 94550, USA

Need to speak English puts burden on Asian scientists

SIR — Masao Ito and Torsten Wiesel point out in Correspondence that factors such as non-standard design or a lack of English on homepages make it harder for the Human Frontier Science Program to identify potential project collaborators or reviewers ("Cultural differences reduce Japanese researchers' visibility on the Web" *Nature* **444,** 817, 2006). However, it is surely not the case that only those whose details are easily found through Web-based searching will be able to compete effectively. Inhomogeneities in the ability of scientists and institutions to perform internationally competitive research have been with us for many decades, and I doubt whether use of the Web has perturbed this pattern to any substantial extent.

I would also dispute Ito and Wiesel's contention that perfecting English-language homepages is a 'simple remedy'. Having to communicate in English to survive, let alone flourish, in the international scientific arena places enormous additional burdens on scientists throughout Asia. Working in a young science-based university in Japan, I see these obstacles being confronted every day. Reading the literature, writing manuscripts, giving oral presentations at conferences — all of these are difficult enough for native speakers of English, but they are fearsome tasks for Asian scientists. In my experience, most English-speaking scientists are blissfully unaware of their magnitude.

The brightest minds in Asian science, adept not only as scientists but also as linguists, will easily possess the wherewithal to create homepages that do them justice. It is with their equally scientifically gifted but less linguistically capable colleagues that we should be concerned.

Ian Smith
Shikanodai-Nishi 3-6-10, Ikoma,
Nara 630-0114, Japan

©2007 Nature Publishing Group

COMMENTARY

Merging and emerging cohorts

How best to study the effects of genes and environment on US health? In the first of two commentaries, **Walter C. Willett** and his co-authors argue that investing in existing studies is the most efficient approach. In the second, **Francis S. Collins** and **Teri A. Manolio** explain their support for a new national cohort.

Not worth the wait

In 2006, the United Kingdom initiated a national long-term health study of 500,000 middle-aged adults that will involve collecting DNA and other biological specimens[1]. Further cohorts are being considered elsewhere in Europe and Asia. Francis Collins (ref. 2 and page 259) has proposed a similar national cohort of several hundred thousand North Americans to enable future studies of the genetic basis of human diseases and individual susceptibility to environmental factors. The cost is estimated to be US$3 billion or more[3].

We are concerned that results from a new cohort would not be available for at least ten years, as five years would be needed for funding, planning and enrolment, and another five for following up even the earliest analyses of the most common diseases; results for most cancers would take longer. We believe that a strategy using existing cohort studies should also be considered. We argue here that this approach can achieve the anticipated objectives of the national cohort more rapidly and more cheaply, and with similar scientific validity.

Already available

Much of what is known about the causes of cancer, heart disease and other illnesses has arisen from epidemiological studies, especially cohort studies. Prospective cohorts, such as the Framingham Heart Study, have had success in identifying key determinants of major diseases, including smoking, physical activity, occupational exposures and adverse effects of medications[4]. In the late 1970s, several much larger cohort studies started to assess lifestyle factors, with many also collecting and storing blood samples for analyses of DNA and plasma[5].

The Table overleaf lists large US cohorts, which total nearly 1,400,000 participants with biologic specimens available for more than 800,000 of them. There are also several ongoing medium-sized cohorts, with typically 5,000–20,000 available biospecimens, which focus on cardiovascular disease[6]. These cohorts also record detailed biological measurements and repeated measures of atherosclerosis.

GETTY

Although most large or medium cohort studies were initially designed to investigate cancer or cardiovascular disease, each routinely ascertains information on other diseases and lifestyle. In the Nurses' Health Study, for example, published endpoints include stroke, diabetes, cataracts, asthma, gallstones, multiple sclerosis, Parkinson's disease, phobic anxiety and quality of life, among many others. The ability to study multiple endpoints within the same population, at modest additional cost, is invaluable in formulating public-health and clinical guidelines, because some environmental exposures may decrease the risk of one disease while increasing the risk of another.

Each existing cohort has tracking systems enabling retrieval of the biospecimens for assay as emergent hypotheses arise. Most cohorts collect comprehensive information on the participants at enrolment, including data on demographic, medical, familial, occupational and lifestyle (such as alcohol intake and diet) characteristics. Such information is usually limited to adult experiences, but lifetime histories are solicited for some exposures such as smoking.

Moreover, participants have already provided informed consent for the use of their biographical data and biospecimens for research. Institutional review boards have approved each study and monitor new ethical issues as they arise. Many investigators already have data-sharing policies in place for collaborations. Reconsent is sometimes needed with existing or new cohorts to accommodate new ethical issues, but this is more easily obtained when a trusting relationship already exists between participants and investigators.

Pooling power

Individual cohorts have already provided key clues to disease aetiology and prevention, but combining data from multiple cohorts offers even greater potential when studying the interplay between environmental and genetic factors. In the past, when findings from individual cohorts seemed inconsistent or inconclusive, data from several cohorts have been combined, sometimes in meta-analyses that attempt to overcome problems of small sample size. But meta-analyses of published reports can also be unsatisfactory because of publication bias towards positive results and because of results that are hard to combine statistically owing to differences in analysis and reporting of data.

Instead, groups of investigators are now collaborating to analyse the primary data from their studies jointly. For example, the Pooling Project of Cohort Studies of Diet and Cancer, which includes more than 25 cohorts from the United States, Europe and Asia with more than 2 million men and women (of which half have biomarkers), has published a series of pooled data on diet and cancer[6,7]. When these data are analysed in a standard manner the findings are highly consistent.

A similar consortium, coordinated by the National Cancer Institute (NCI), is systematically genotyping samples and combining results for breast and prostate cancer to investigate gene–environment and gene–gene

©2007 Nature Publishing Group

interactions[8,9]. So far, this consortium includes about 10,000 cases of both cancers, along with matched controls from the same cohorts, which allows for robust exploration of candidate genetic pathways. To attain these same numbers with a US national cohort would require following up 500,000 women for 20 years and 500,000 men for 10 years — or 15 million person-years of follow-up.

As an example of the NCI consortium's ability to address new leads rapidly, within weeks of the publication of a polymorphism for prostate cancer risk[10], the consortium was able to confirm this finding in seven individual cohorts. Indeed, whole-genome scanning of 10,000 cancer cases and 10,000 controls could identify almost all common genetic variants even weakly associated with breast and prostate cancer at an estimated cost of $5–10 million.

Existing large US cohort studies with biospecimen repositories containing blood and/or DNA samples

Study	Year biospecimen collection began	Total cohort size	No. with stored biological samples
Health Professionals Follow-Up Study	1986	52,000	30,000
Nurses' Health Study I	1989	122,000	63,000
Washington County Study	1989	33,000	33,000
Women's Health Study	1992	40,000	28,000
Women's Health Initiative	1993	162,000	162,000
NCI PLCO Study	1994	155,000	70,000
Nurses' Health Study II	1996	116,000	60,000
American Cancer Society CPS-II LifeLink Study	1998	184,000	109,000
Multiethnic Cohort Study	1996	215,000	80,000*
Vitamins and Lifestyle (VITAL) Cohort	1999	78,000	54,000
Agricultural Health Study	1999	90,000	35,000
Southern Community Cohort Study	2002	90,000*	80,000*
Black Women's Cohort Study	2005	59,000	41,000*
Total	**-**	**1,396,000***	**845,000**

*Expected totals upon completion

Dispelling myths

Certain limitations of the existing cohorts have been raised[2,11]. One oft-repeated concern is that existing cohorts are not representative of the US population. However, representativeness is not needed to determine the relation between genetic or environmental factors and disease risk; for example, the effects of smoking were first revealed in a study of British doctors. In many cases, cohort studies are purposefully nonrepresentative to maximize the quality of data or to emphasize the contrast between environmental exposures. A cohort of Seventh-Day Adventists, for example, allows the investigation of high consumption of soya products, and studies of nurses or doctors provides insight into long-term behaviours thought to be healthy but under-represented in the general population.

A related myth is that population subgroups, whether racial, gender, or religious, must be represented in proportion to their prevalence in the general population in any single study to be confident that the results apply to these subgroups. But what is really needed is a subgroup large enough to examine the exposure and disease association within that subgroup. The cohorts with biospecimen repositories listed in the Table include large numbers of African-Americans, Hispanics and other minority groups. In our view, the US National Institutes of Health (NIH) research portfolio already has adequate statistical power to investigate genetic variations and environmental factors within most major ethnic subgroups except for Asians.

Another criticism of existing cohorts is that self-reported measures of diet, physical activity, and other environmental exposures are less useful than more time-intensive and costly objective measurements. But because diet and physical activity vary substantially over time for most individuals, even an objective measure of short-term exposure (which rarely exists) can be inferior to a questionnaire that solicits less precise information over a longer period. Also, biomarkers using plasma, red blood cells, DNA and nail samples are now being used to assess many dietary and other environmental factors, and can be integrated into existing cohorts. For many biomarkers, the existing specimen repositories will be invaluable because millions of person-years of follow-up are already available.

Layered approach

Because most existing cohorts enrolled middle-aged or older adults, information is sparse on childhood and early adult years. Young age may be a critical window of exposure to cancer risks, for example[12]. Ultimately, cohorts established during childhood will be more informative, but decades of follow-up will be required for information on adult diseases. Although not perfect, some existing cohorts seek information on early-life exposures based on interviews with the adult participants.

In addition, the Nurses' Health Study II has enrolled around 25,000 offspring of existing participants between the ages of 10 and 14; some have already been followed up for ten years. Biological specimens are not yet available, in part because of the cost and complexities regarding informed consent from children, but DNA could be collected at age 21. This cohort is probably large enough to study most common outcomes, say type 2 diabetes or asthma, but not less-common outcomes such as specific cancers. We believe a layered approach makes sense because most common outcomes, such as hypertension, can be addressed by detailed studies of medium-size cohorts, and less common outcomes, such as cancer, can be studied more cheaply with less-intensive approaches using large cohorts.

Substantial investments have already been made, mainly by the NIH, to create cohorts with more participants than the proposed new cohort, and sufficient follow-up to conduct powerful genetic analyses for major cancers and other health endpoints. Although the existing cohorts have some gaps in data and specimens, these can be filled with relatively modest investments. In an era of limited research funding, a wise strategy would be first to use existing resources more efficiently before embarking on an extraordinarily expensive new cohort, which will provide little or no return for the next decade or more. ■

Walter C. Willett is in the Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts 02115, USA; his co-authors are William J. Blot, Graham A. Colditz, Aaron R. Folsom, Brian E. Henderson and Meir J. Stampfer. For additional information, e-mail walter.willett@channing.harvard.edu.

1. Watts, G. *BMJ* **332,** 1052 (2006).
2. Collins, F. S. *Nature* **429,** 475–477 (2004).
3. Spivey, A. *Environ. Health Perspect.* **114,** A466–A467 (2006).
4. Samet, J. & Munoz, A. *Epidemiol. Rev.* **20,** 1–14 (1998).
5. Langholz, B., Rothman, N., Wacholder, S. & Thomas, D. C. *Monogr. Natl Cancer Inst.* **26,** 39–42 (1999).
6. National Heart, Lung and Blood Institute population studies database. http://apps.nhlbi.nih.gov/popstudies/ (accessed 24 August 2006).
7. Smith-Warner, S. A. *et al. Am. J. Epidemiol.* **163,** 1053–1064 (2006).
8. Hunter, D. J. *et al. Nature Rev. Cancer* **5,** 977–985 (2005).
9. Feigelson, H. S. *et al. Cancer Res.* **66,** 2468–2475 (2006).
10. Amundadottir, L. T. *et al. Nature Genet.* **38,** 652–658 (published online 7 May 2006).
11. Secretary's Advisory Committee on Genetics, Health and Society. http://www4.od.nih.gov/oba/SACGHS/public_comments.htm (May 2006).
12. Land, C. E. *et al. Radiat. Res.* **160,** 701–717 (2003).

Acknowledgment The authors acknowledge the helpful suggestions of D. Hunter.

©2007 Nature Publishing Group

Necessary but not sufficient

The proposal advocated in the preceding Commentary by Willett *et al.*[1], namely to extend existing cohort studies rather than start a new large-scale prospective study from scratch, has many merits. Indeed, a US National Institutes of Health (NIH) study group that assessed the pros and cons of various models in 2004 considered this option in some depth, and their report[2] made many of the same points.

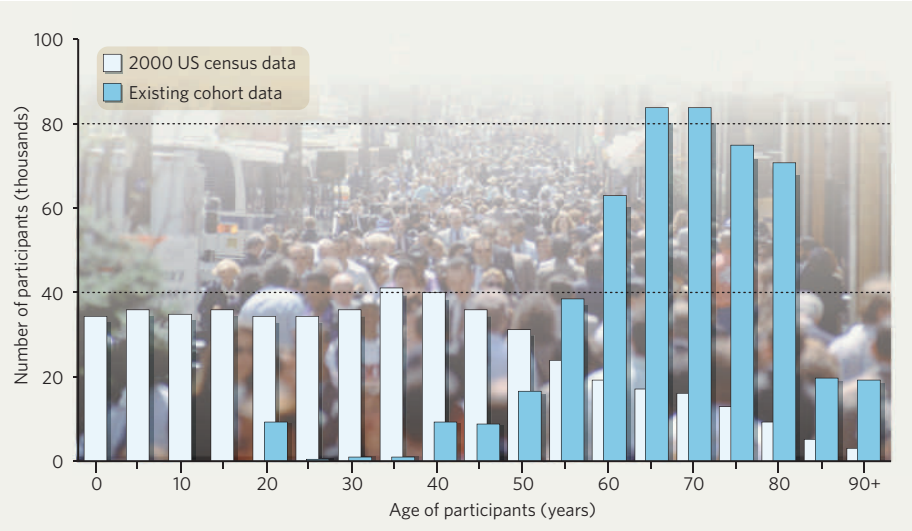
Certainly, assembling existing cohorts into a large consortium would provide a powerful resource for investigating genetic and environmental factors in health and disease. The argument that this method is likely to be less costly than a new cohort, and would yield results more quickly, carry considerable weight. But Willett *et al.* do not address all of the suboptimal aspects of this approach. Those should be clearly noted, lest expectations of such a consortium exceed what it is likely to deliver.

First, there is the issue of standardization. Phenotypic measures used by the existing cohorts, although standardized within cohorts, have not followed uniform procedures across studies, and so there will be significant challenges to merging data from different studies in a valid way. Moreover, key environmental exposures or risk factors will almost certainly differ systematically across cohorts. Combining studies that were focused on specific population subgroups will therefore introduce biases that can be corrected only by limiting the analysis to the lowest common denominator of valid, unbiased exposures.

Second, the reliance on legacy studies fails to take advantage of new tools for measuring dietary intake, physical activity and environmental exposures (as are now being supported through the NIH Genes and Environment Initiative[3]), because many of these measurements — such as precise ambulatory data — cannot be made on stored biospecimens.

Third, representation has been a major concern driving the national cohort proposal. Despite recent attempts to improve representation of minorities and socioeconomically disadvantaged participants in newer cohorts, the proportions are still far below their representation in the US population. There is also substantial under-representation among men and participants from the south, as well as those with lower levels of education, although these might be addressed somewhat by statistical adjustments.

Fourth, under-representation of people younger than the age of 50 is substantial in these existing cohorts (see Figure) and will only get worse with time. If we wish to address complex disease risk across lifespan, we need to study diseases developing in adolescence and young adulthood, such as asthma, autoimmune disease and major psychoses. Even the investigation of mid-life diseases would be limited by lack of stored biospecimens in earlier life.



GETTY

Comparison of an estimated distribution of a 500,000-person cohort based on existing cohort data[2], with the 2000 US census.

Finally, the full value of a large-scale cohort study will depend on free and open access to the data by all qualified investigators. This may be difficult to achieve with a combination of existing cohorts, given the expectations of current investigators about control of the data, and consent limitations by existing study participants.

More ways than one

There is no question that a new cohort study would require many years to implement and to generate results, although useful findings would be available on more common diseases within five years of cohort recruitment[4]. We agree with Willett *et al.*, therefore, that it is reasonable in the interim to seek ways to form consortia of existing studies. But these two models need not be thought of as mutually exclusive.

We must also recognize that environmental exposures (including emerging infections) and preventive or therapeutic interventions will probably change dramatically in the next two decades. Limiting our research enterprise to the exclusive study of existing cohorts, especially without collection of new risk information and recruitment of participants under-represented in existing studies, may ultimately jeopardize our ability to address these evolving health risks in an epidemiologically rigorous manner.

Admittedly, this discussion remains hypothetical, because serious budgetary challenges make a new national cohort an unlikely prospect at the present time. Some may wonder whether the United States can afford both an expansion of existing cohorts and a new national cohort. We believe the real question is whether it can afford not to do both, given the enormous and growing healthcare costs of complex diseases. Finding the genetic causes of even one of these diseases could potentially save billions of dollars in medical costs if appropriate preventive interventions can be developed. Despite the fiscal realities, therefore, we must continue to make the case both for a merging of cohorts now, and the founding of a more rigorously designed national cohort in the future when funds are available. Although recognizing the massive uncertainties in budget situations and priorities, we believe that future generations will wonder why we didn't try as hard as possible to get both of these kinds of studies underway. ■

Francis S. Collins and Teri A. Manolio are at the National Human Genome Research Institute, National Institutes of Health, 31 Center Drive, Bethesda, Maryland 20892-2152, USA.

1. Willett, W. C. *et al. Nature* **445,** 257–258 (2007).
2. www.genome.gov/Pages/About/OD/ReportsPublications/PotentialUSCohort.pdf
3. NIH Genes and Environment Initiative www.gei.nih.gov/
4. Manolio, T. A., Bailey-Wilson, J. E. & Collins, F. S. *Nature Rev. Genet.* **7,** 812–820 (2006).

Acknowledgements We acknowledge A. Guttmacher, E. Harris and L. Rodriguez for their advice.

©2007 Nature Publishing Group

BOOKS & ARTS

Social outcasts

The most commonly studied social insects are the Hymenoptera, but what about all the rest?

The Other Insect Societies
by James T. Costa
Belknap Press: 2006. 812 pp.
$59.95, £38.95

Jenai M. Milliser and George W. Uetz

For much of its history, the study of insect sociality has been dominated by the study of ants, bees and wasps (the order Hymenoptera) and driven by a single paradigm (genetic relatedness and kin selection). The other social arthropods have been largely ignored or relegated to fringe status. James T. Costa seeks to correct this oversight with his book *The Other Insect Societies*. He succeeds in showcasing the true diversity of social behaviour among insects, spiders and crustaceans, and provides support for a set of alternative hypotheses about social evolution that should stimulate research and fuel scientific debate for years to come.

Perhaps it is not surprising that bees, ants and wasps have historically taken the limelight at the expense of the 'other insect societies'. They are nearly ubiquitous and economically important, and have fascinating social organization and seemingly inexplicable altruistic tendencies. As a result, these hymenopterans provide a well-researched database on which to develop an overarching framework of evolutionary theory. W. D. Hamilton's theory of kin selection (*J. Theor. Biol.* **7**, 1–16, 17–52; 1964) explains the role of genetic relatedness in the evolution of altruistic behaviour, such as reproductive division of labour, with sterile members effectively caring for the reproductive members. This has become a pivotal concept in arthropod sociobiology, popularized in E. O. Wilson's book *The Insect Societies* (Belknap, 1971). However, as Costa's impressive compilation makes clear, there are likely to be many evolutionary pathways to sociality that may or may not involve kin selection.

Looking beyond kin selection and relatedness, Costa emphasizes the need to disentangle the concepts of eusociality and sociality, and to abandon the widely embraced 'terminological hierarchies' of sociobiology (subsocial, communal, quasisocial, semisocial, eusocial). The path delineated by this terminology is ultimately framed around the key criteria that define eusociality: overlapping generations, reproductive division of labour, and sterile worker castes. Set forth most prominently by Wilson in *Sociobiology* (Belknap, 1975), these criteria for eusociality have been the gold standard by which all non-eusocial arthropod societies have been measured and compared. Costa effectively argues that this terminology has so constrained our perspective that the role of ecological factors in the evolution of arthropod sociality has been largely unrecognized. That is not to say that the classic criteria of eusociality are without value; however, when used as a framework to examine non-eusocial insect societies, they are often as restrictive as they are enlightening.

J. T. COSTA

Let's stick together: weevil larvae (of the order Coleoptera) aggregate on a leaf.

Ultimately, Costa synthesizes ideas from his own work with those of others to create new categorizations and terminologies with broader applicability. Constructing his review of insect societies around four primary categories — maternal and biparental care, paternal care, 'fortress defenders' and 'larval societies' — he applies a framework that concentrates as much on the ecological context of sociality as the type of social structure.

The book provides a careful treatment of both the well-studied and the relatively obscure taxa that make up the non-eusocial social insects. Several chapters are focused on selected representatives of the Coleoptera (for example, bark and ambrosia beetles, weevils, fungus beetles, whirligig and burying beetles), Lepidoptera (social caterpillars, webworms, swallowtails and silk moths) and Hemiptera (aphids, treehoppers, cicadas and leafhoppers) and are organized by ecological similarities. Costa also delves into obscure resources to provide fascinating accounts of the lesser-known taxa with relatively scarce ecological and behavioural data, such as webspinners (Embiidina), sawflies (Symphyta), earwigs (Dermaptera) and thrips (Thysanoptera). He also includes non-insect arthropods, such as arachnids, millipedes, centipedes and crustaceans.

Although others have attempted to synthesize the literature on non-eusocial insect

©2007 Nature Publishing Group

societies, this is the first major single-author monograph on the topic for more than a decade. We found little to criticize, even in such a large book. Perhaps graphical representations of data might have improved the strength of certain scientific arguments, in place of some of the line drawings, which seem oddly placed and in many cases have no clear connection to the text. On the plus side, many of the taxa covered in the book are illustrated with stunning colour photographs.

Costa takes a refreshingly unbiased approach to examining these fascinating societies, providing insight into how non-eusocial insect societies are structured, as well as details of their systematics, natural history, ecology and anatomy. His writing is scholarly enough to appeal to professionals, yet it is accessible enough to enlighten any interested reader. His book gives the 'other insect societies' well-deserved time in the spotlight, and will be a valuable and appealing addition to the libraries of academics and readers of scientific natural history alike. ■

Jenai Milliser and George Uetz are in the Department of Biological Sciences, University of Cincinnati, Cincinnati, Ohio 45221, USA.

Beyond fractals

Superfractals: Patterns of Nature
by Michael F. Barnsley
Cambridge University Press: 2006.
464 pp. £25, $35

J. R. Mureika

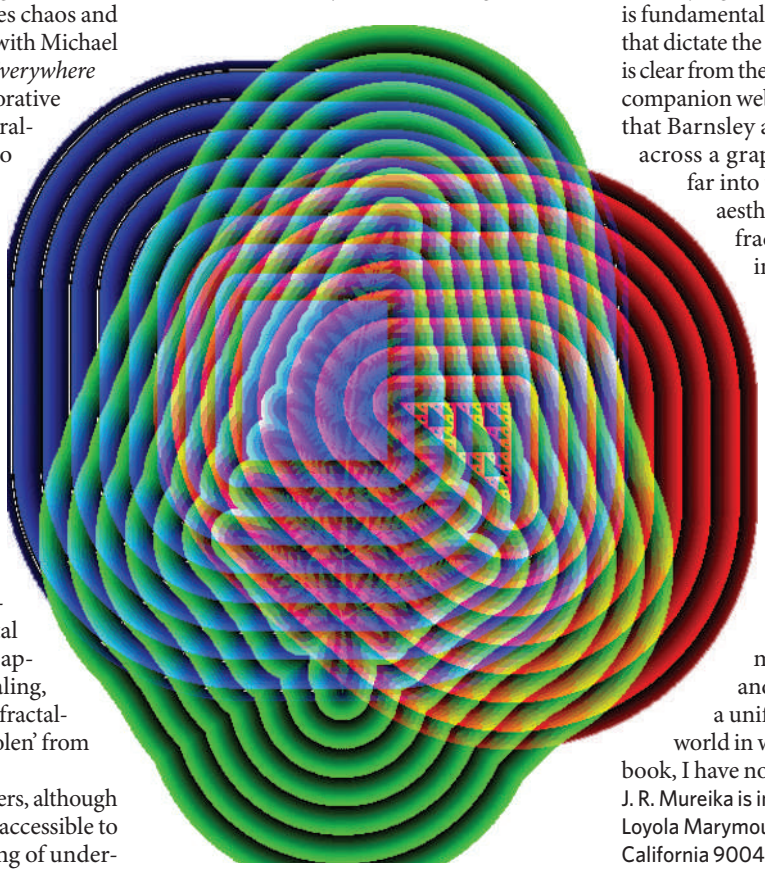
Since Benoît Mandelbrot first brought fractals to the attention of the world, their self-similar mystique has wormed its way into the physical, biological and actuarial sciences, as well as psychology and abstract art. Describing the seemingly disordered world with such geometrical simplicity has long been a target of reductionism. However, using simple fractals hasn't been entirely successful: the world isn't quite that easy to describe with algebra.

Anyone whose research involves chaos and recursion will already be familiar with Michael Barnsley's seminal work *Fractals Everywhere* (Academic Press, 1988). A collaborative effort with researchers at the Australian National University has led to the development of superfractals, allowing Barnsley to offer a new link between the physical realities of the Universe and the symmetries of mathematics.

Briefly, a superfractal is the result of an iterative application of randomly selected but well-defined transformations to a picture — a hybrid recursion that results in a wide range of beautifully intricate structures. *Superfractals* colourfully conveys to the reader this next step in chaotic harmony, highlighting the unification of information theory and physics. New to the fractal arena is the use of code space mappings, fractal tops and colour stealing, in which the pixels of the resulting fractal-like attractor are assigned hues 'stolen' from a base image.

The book is not easy for lay readers, although the author states that it should be accessible to anyone with a basic understanding of undergraduate mathematics. The overall structure of the text is well thought out, and at the outset of each chapter the author delineates exactly which salient issues will be addressed in the following pages, making for a good executive summary. But because the book is very linear in its presentation, with subsequent chapters relying strongly on results derived in the previous ones, any readers who 'dive in' part way through will have difficulty gaining a clear understanding of Barnsley's conclusions.

Despite this, the narrative unfolds in an almost jovial manner, allowing the author's contagious passion for the subject to shine through. Starting from mathematical first principles, Barnsley leads us through the evolution of algebraic sophistication and on to new and wondrous applications of the 'chaos game' and iterative system functions. Also known as a Markov chain Monte Carlo algorithm, this scheme provides a connection to basic fractals, such as the Sierpinski triangle. More importantly, the author shows how randomness might play a crucial role both in determining the rich complexity in nature and in the revolutionary field of bioinformatics.



The interplay between abstraction and the 'real world' is accented by Barnsley's set-theoretic formalism, but this makes it surprisingly easy to understand the digitization techniques, image transformations, and formation of self-similar fractal features. Computer-generated superfractals are interlaced with the equally artistic mathematical body of theorems, proofs and functional mechanisms that define the beauty within the self-similarity. These images encapsulate the transformations being described by the mathematics, exemplifying the fractal dynamics at work without any need for the technical jargon. In a sense, the reader is treated to two books at once, thanks to Barnsley's union of art and mathematics.

The occurrence of superfractal structure in the real world is highlighted by the author's choice of figures. The functions and graphs one might expect to find in a mathematically intensive treatise are nowhere to be seen. Instead, the pages are adorned with colourful pictures of leaves and animals, and there are even some photographs. This relaxed use of graphics is no accident; rather, it is a consequence of the underlying theme of the book, that the world is fundamentally described by the simple rules that dictate the chaotic overtures of iteration. It is clear from these images, as well as those on the companion website (www.superfractals.com), that Barnsley and colleagues have stumbled across a graphical technique that reaches far into the burgeoning field of fractal aesthetics. The applications of superfractals in the computer graphics industry are obvious, and stand to transform the way in which textures and natural patterns are mimicked.

Overall, *Superfractals* would be a superb addition to the bookshelves of any scientists who use fractal analysis techniques in their research, be they physicist, biologist or economist. The author concludes by promising that the introduction of superfractals will revolutionize the way mathematics, physics, biology and art are combined, to produce a unified description of the complex world in which we live. After reading this book, I have no doubt that he is correct. ■

J. R. Mureika is in the Department of Physics, Loyola Marymount University, Los Angeles, California 90045, USA.

©2007 Nature Publishing Group

Heady days in Princeton

Pursuit of Genius: Flexner, Einstein, and the Early Faculty at the Institute for Advanced Study
by Steve Batterson
A. K. Peters: 2006. 314 pp. $39

John Stachel

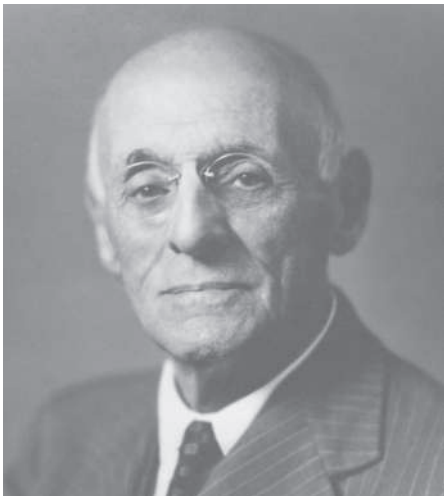
Punching "advanced study" into Google's search engine brings up more than 2 million entries, but heading the list is the Institute for Advanced Study (IAS) in Princeton. Born of the union between private philanthropy and individual vision, the institute was a typically American marriage consciously modelled on the one that gave birth in 1876 to Johns Hopkins University, the first research university in the United States.

The IAS was itself the model for many other centres, including the Mexican Centro de Investigación y de Estudios Avanzados (CINVESTAV), set up 45 years ago, which was typical of the way research on a national scale is sponsored and financed by national governments. (I have spent time at both the IAS and CINVESTAV.) The United Nations University's Institute of Advanced Studies, set up just a decade ago, suggests a possible future pattern, with research being envisaged, controlled and conducted by international agencies on an international scale.

It is interesting to read in *Pursuit of Genius* by Steve Batterson how the IAS fared in the early days under the direction of its founding genius, Abraham Flexner. Flexner made his name with a report in 1910 on the quality of medical education in the United States. His brother Simon, an MD, had persuaded John D. Rockefeller to let him run a medical institute on the Johns Hopkins model; it also employed Flexner for some time. Meanwhile, two wealthy Bamberger siblings, Louis and Caroline, sold their Newark department store just before the 1929 economic crash and wanted to found a medical school. They approached Flexner, but he convinced them to abandon the idea in favour of an institute of higher learning, a hybrid centre with Johns Hopkins' scope and the Rockefeller Institute's methods.

Flexner's first great success, chronicled in detail by Batterson, was the founding of the School of Mathematics, in which he had the unintended help of Adolf Hitler. Einstein had originally agreed to spend half his time at the new institute, but when the Nazis came to power in Germany in 1933, he realized he could not return to Berlin. He was joined by mathematicians Hermann Weyl (after much soul searching and a nervous breakdown) and Kurt Gödel (who see-sawed between Vienna and Princeton) at the new institute's first home in Fine Hall, the Princeton mathematics building.

There were other successes during the decade, such as the building of a separate campus for the institute and the founding of several



IAS ARCHIVES

Abraham Flexner helped to establish the Institute for Advanced Study at Princeton.

other schools. But growing friction between Flexner's autocratic methods and the faculty's desire to participate in the institute's governance led to a struggle that culminated in Flexner's resignation in 1939. But the power struggle continued, as Batterson makes clear in a rather perfunctory final chapter giving a high-speed sketch of the intervening sixty-odd years of the institute's history.

The book is based primarily on a meticulous study of the institute's documentary records and the papers of some key players in its early history. Batterson tells a fascinating story, largely unvarnished, including such unsavoury details as the anti-Semitic atmosphere of Princeton at that time. The institute's patrons, its director and much of its initial faculty were Jewish, yet Flexner repeatedly refused to face up to anti-Semitism in the US educational system, let alone do anything to fight it.

More amusing are Batterson's detailed accounts of the turf war between Princeton University, particularly the mathematics department headed by Luther Eisenhart, and the institute's budding mathematics section, headed by Oswald Veblen.

The emphasis on documentary detail does not always generate a feeling for the setting in which these dramas took place. Einstein's letter to Belgium's Queen Elizabeth, written soon after his arrival, better conveys a sense of how incredibly provincial Princeton seemed to someone coming from Weimar Berlin: "Princeton is a wonderful little spot, a quaint and ceremonious little village of demigods on stilts… Here the people who compose what is called 'society' enjoy even less freedom than their counterparts in Europe. Yet, they seem unaware of this restriction since their way of life tends to inhibit personality development from childhood."

The use of Einstein's name in the title seems to be a hook to grab readers. Although he does play an important role in the story, it is no greater than that of many other early faculty, such as Veblen. The account of Einstein's private life is not always accurate and provides fewer new insights than the more fascinating account of Weyl's; those who want a more intimate picture should read Jamie Sayen's *Einstein in America* (Crown, 1985). Readers inclined to accept the myth that Einstein was completely isolated after his move to Princeton will be surprised at the large public role he played in shaping the fate of the institute, particularly in finally bringing down Flexner. ■

John Stachel is emeritus professor of physics and director of the Center for Einstein Studies, Boston University, Boston, Massachusetts 02215, USA.

Display elements

The Periodic Table: Its Story And Its Significance
Eric M. Scerri
Oxford University Press: 2006. 368 pp. £19.99, $35

Bernadette Bensaude-Vincent

The Periodic Table by Eric Scerri is the first major English-language monograph on the history of the periodic system since Jan van Spronsen's *The Periodic System of Chemical Elements* (Elsevier, 1969). Like van Spronsen, Scerri tries to understand the construction of the periodic table within the context of the history of matter theories. And like van Spronsen, he insists that it results from the collective work of many chemists. Chapter 3, which begins with the early triads derived from Prout's hypothesis and presents six co-discoverers of the periodic system, provides a summary of van Spronsen's classic study. However, his superficial overview of successive notions of elements mainly reinforces some popular clichés.

At this point, Scerri departs from van Spronsen. The historical sketch just provides the background for an exposition of his philosophical views on chemistry. Scerri, the editor of the journal *Foundations of Chemistry*, is interested above all in the question of whether chemistry can be reduced to quantum mechanics. He resolutely opposes Paul Dirac's 1929 claim that "the underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are completely known". The main purpose of his book is to

©2007 Nature Publishing Group

demonstrate that the periodic system has not been reduced to physics and cannot be deduced from quantum theory.

Scerri pays attention too to the philosophical choices underlying the construction of the periodic table. The chapter on Dmitry Mendeleev's process of discovery emphasizes that Mendeleev had an abstract 'metaphysical' notion of chemical elements that he clearly distinguished from Antoine Lavoisier's view of elements as simple, concrete substances — a conceptual shift first emphasized some 20 years ago. In addition, Scerri has misunderstood the epistemological status of Mendeleev's abstract notion of the element. Far from reviving a metaphysical notion, Mendeleev did his best to promote a positive, if abstract, notion of the element. Not only did he characterize this invisible entity by an individual quantitative property — its atomic weight — but he defined it by analogy with other basic concepts of chemistry, stating that the distinction between the element and simple substance was like the distinction between atoms and molecules.

Scerri is more original in his detailed account of Mendeleev's famous predictions of unknown elements. He revises the usual success story in pointing out that Mendeleev failed as often as he succeeded. He thus discusses a question first raised by Stephen Brush: whether predictions were a decisive factor in the acceptance of the periodic system. Scerri emphasizes the importance of Mendeleev's ability to fit all the elements into the system.

The core of Scerri's argument is to be found in the chapters dealing with the evolution of the periodic system with regard to changes in atomic theory in the aftermath of Mendeleev's discovery. Here Scerri makes a plea for the autonomy of chemistry. He convincingly argues that the abstract notion of the element was crucial to rescuing the periodic system in the light of the discovery of isotopes. He also rightly notes that Niels Bohr's atomic model relied heavily on spectroscopic data, rather than on theoretical calculations. Against repeated claims that chemistry has been reduced to physics, it is always useful to keep in mind that early quantum physics was based on chemical data.

Finally, in considering a variety of visual representations of the periodic system, Scerri advocates a system known as Charles Janet's left-step table, because it rests on the concept of elements as basic substances, rather than on physical properties.

All practitioners of chemistry, from researchers and teachers to engineers, seem have an opinion about what the periodic table should look like, and many of them continue Mendeleev's work in an attempt to propose better graphic representations. Despite the standard format recommended by the International Union of Pure and Applied Chemistry in 1985, with groups numbered from 1 to 18, new tables are invented every year. In 1973, Edward Mazurs reviewed hundreds of visual representations of the table and distinguished 146 structural types. The periodic system, then, is like a monument, forever inviting new creative designs. ■

Bernadette Bensaude-Vincent is in the Department of Philosophy, Université Paris X, 200 avenue de la république, 92001 Nanterre, France.

SCIENCE IN CULTURE

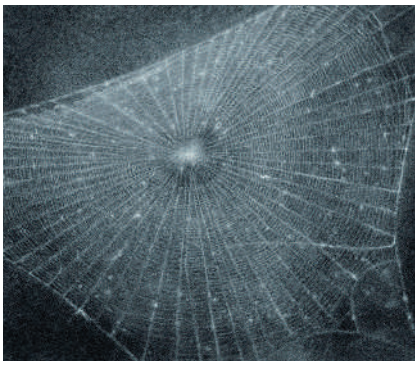
Drawn to nature

Vija Celmins' graphite and charcoal drawings are inspired by the natural world.

Colin Martin

Drawing is central to the work of Latvian-born US artist Vija Celmins. Her haunting, monochrome depictions of limitless expanses of ocean, nocturnal skies and deserts lack a point of reference, such as the horizon, or a depth of field. Usually copied meticulously from photographs she has taken herself, each image takes months to draw. 'Vija Celmins: A Drawings Retrospective', an exhibition originally shown at the Centre Georges Pompidou in Paris and on view at the Hammer Museum in Los Angeles, California, from 28 January (www.hammer.ucla.edu/exhibitions/119), brings together 68 of Celmins' drawings made over a 40-year period.

V. CELMINS/ADAGP, PARIS



V. CELMINS/ADAGP, PARIS

In 1968, Celmins began using photographs of outer space to translate the idea of the night sky into an abstract composition, an approach she then adopted in her ocean drawings. "The sight of the waves miles out, their dutiful and frenetic solitude, their dull indifference to their fate," mused novelist Colm Tóibín in his catalogue essay.

Celmins drew *Untitled (Big Sea #1)* (shown above, left image) in 1969 using graphite on a sheet of paper covered with a light grey acrylic ground, a technique she favours to avoid digging her pencil into paper. By leaving a thin grey border around her images (not shown here), Celmins signals that her subjects are clearly defined objects and not fragmentary images. The subject matter is clearly her photograph, not an actual seascape. "I don't imagine the ocean and try to recreate a memory of it when I'm doing the art," Celmins said in 1992. "I explore a surface through drawing it. The image gets controlled, compressed and transformed."

Over time, Celmins' ocean drawings became denser and she created perspective by using thicker layers of graphite to darken the lower parts of her drawings. Repetition became more significant in her work, and she began producing series of similar images.

In 1983, thinking she had exhausted drawing as her medium, Celmins returned to painting. But she began drawing again in 1994, using a different technique. She applied charcoal by hand directly onto paper, and used erasers to rub through the charcoal to create images by exposing the white surface of the paper.

The subjects of Celmins' latest series of drawings are delicate, light spider webs, shown against dark charcoal backgrounds. *Web #1* (shown above, right) from 1998 conveys the translucent quality of a web in an image that has a sense of discovery and wonder. The series was inspired by a scientific publication, James Henry Emerton's 1902 book *The Common Spiders of the United States*. "If I wasn't an artist, I think I would have liked being a scientist," said Celmins.

Colin Martin is a London-based writer.

©2007 Nature Publishing Group

GALAXY MORPHOLOGY

Out of order

Recent observations of vast numbers of galaxies may pose problems for the galaxy classification scheme proposed by Edwin Hubble.

Sidney van den Bergh

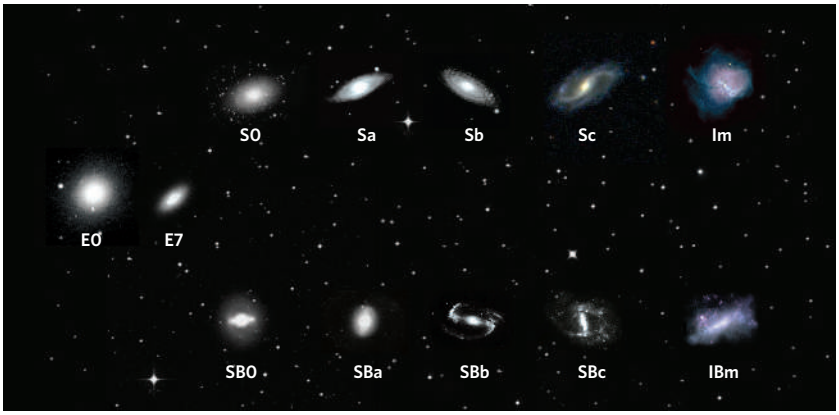
Galaxies are like people: the better you get to know them, the more peculiar they often seem. Classifying and explaining the appearance of objects that occur in such an array of shapes, masses and sizes is a major challenge. As astronomer Allan Sandage wrote recently: "The danger in devising a classification from scratch is that, if the classifier has preconceived notions of what those explanations might be, then the classification becomes circular if it is later used to provide the explanation." To complicate matters, the Hubble Space Telescope (HST) has shown that the appearance of galaxies, like that of people, changes systematically with age.

Modern galaxy classification systems evolved from the 1926 work of Edwin Hubble. He placed galaxies on a linear sequence E0–E7–Sa–Sb–Sc, with ellipticals designated as E and spiral galaxies S. In this sequence, E0s are round ellipticals; E7s are highly flattened ellipticals; Sa galaxies have large elliptical-like cores surrounded by smooth, tightly coiled spiral arms; Sb galaxies have smaller cores and more open arms; and Sc galaxies have small cores that are surrounded by very open, patchy, spiral arms. Hubble also noted a dichotomy between normal and barred spirals (as shown in the figure). He labelled galaxies that did not fit comfortably into his classification scheme 'peculiar'.

We now know that there are two classes of peculiar object: normal galaxies that have been catastrophically reshaped by violent tidal encounters or mergers, and irregular galaxies that contain bright young stars and no nuclei, and show no regular spiral structure. Unexpectedly, the normal and barred counterparts Hubble observed for spirals extend to irregular galaxies (such as Megellanic irregulars (Im and IBm); see figure).

Ten years after formulating his classification system, Hubble suspected that an additional class — which he dubbed S0 galaxies — might need to be interposed between spirals and ellipticals. This idea was strengthened by observations made by his successors during the next two decades. However, the fit of S0s into Hubble's tuning-fork diagram has always been uncomfortable. They tend to be fainter than either elliptical or Sa galaxies, not intermediate between them as might be expected.

Subsequent work has only complicated the picture. HST observations have shown that the morphology of galaxies changes systematically with look-back time. Like humans, young galaxies develop fast. They gain mass from mergers with lesser galaxies and from inflow of gas, then grow old slowly as their structure relaxes and their stellar content ages and evolves. HST images have shown that elliptical and Sa galaxies settle into a stable morphology very early on, whereas Sc galaxies take much longer.



Hubble's 'tuning-fork' diagram of galactic classes, including normal and barred counterparts.

Luminous nearby Sc galaxies have well-ordered spiral arms, whereas distant ones show chaotic spiral structure. For look-back times of up to five billion years, spirals look like standard nearby objects in HST images. At greater light travel times they appear more and more peculiar. And, at redshifts corresponding to a look-back time of about ten billion years it becomes difficult or impossible to shoehorn them into the Hubble tuning-fork diagram. So, the Hubble classification scheme breaks down entirely for galaxies at distances greater than ten billion light years.

Taken at face value, the Hubble classification scheme suggests that galaxies (with the possible exception of S0s) lie on a single continuous sequence ranging from E through Sa, Sb, Sc and Im. But this seems to conflict with modern, high-precision observations of the luminosities and colours of enormous samples containing tens of thousands of galaxies. Such observations strongly suggest that galaxies fall into two distinct zones in a colour-luminosity diagram. It is not clear how the apparent dichotomy of galaxy characteristics, seen in the distribution of galaxies over the colour-luminosity diagram, can be reconciled with the continuous change of galaxy characteristics along the Hubble classification sequence.

Astronomers generally agree that the Hubble sequence results from changes in population mix, with ellipticals containing mostly old stars and Im galaxies comprising mainly young ones. But the apparent dichotomy shown by galaxies when classified by colour and luminosity is attributed to the different properties of 'high-' and 'low-mass' galaxies. The apparent dichotomy between the sequence of high- and low-mass galaxies arises because shock heating of gas in galaxies more massive than $10^{12}\,M_{\odot}$ abruptly halts star formation, whereas cooling and star formation can continue in galaxies with lower masses. A grand unifying scheme that incorporates both the continuity of the Hubble diagram and the dichotomy in the galaxian colour-magnitude diagram does not yet seem to be in sight.

Albert Einstein and many others have commented on the effectiveness of mathematics for formulation of the laws of nature. As a result, science sometimes evolves in those directions in which mathematics can be applied. However, several areas, including friction, turbulence and morphological classification, remain largely in the mathematical wilderness. Progress in galaxy morphology has mainly resulted from the remarkable human capacity to recognize patterns. This suggests that a deeper understanding of the nature of the differences between galaxies in the colour-luminosity diagram may come from careful inspection of the images of the brightest ones. ■

Sidney van den Bergh is at the Dominion Astrophysical Observatory, Herzberg Institute of Astrophysics, National Research Council, 5071 West Saanich Road, Victoria, British Columbia, V9E 2E7, Canada.

CONCEPTS

©2007 Nature Publishing Group

NEWS & VIEWS

INFLUENZA

Fatal immunity and the 1918 virus

Yueh-Ming Loo and Michael Gale Jr

Innate immune defences are our first line of protection against infection by viruses and are essential in limiting viral disease. But their reaction to the 1918 influenza virus could have been deadly.

The devastating Spanish influenza A virus infected around a third of the world's population during the pandemic of 1918. With mortality rates more than 25 times that of other influenza pandemics, the 1918 virus killed over 40 million people worldwide[1]. Influenza A virus is transmitted among humans and domestic animals, and when different strains infect the same cell, mixing of the two viral genomes (or 'reassortment') can generate new strains with epidemic or pandemic potential.

Studies of genetically engineered influenza A viruses containing some or all of the genes from the 1918 virus suggest that this virus might have acquired a unique combination of genes that caused the disease's unusual severity[1,2]. However, the specific actions of these genes and their contribution to the virulence of the 1918 pandemic have remained elusive. On page 319 of this issue, Kobasa *et al.*[3] provide the first report of the effects of the reconstructed 1918 influenza virus in monkeys. Their findings link the unprecedented lethality of the 1918 pandemic virus to an aberrant innate immune response.

During infection, an invading virus is recognized by specialized cellular proteins that engage viral nucleic acids or proteins and trigger signalling pathways within the infected cell. These pathways culminate in the production of molecules of the innate immune system — our first line of protection against infection — called α/β-interferons, proinflammatory cytokines and chemokines[4]. This group of proteins includes interleukins, which are essential for activating immune cells, and chemoattractant proteins, which recruit immune cells to sites of infection. The innate immune response is triggered over a course of minutes to hours after infection, and these secreted proteins act collectively, both locally and systemically, to direct the expression of a wide range of proteins that act directly against the virus or stimulate inflammation. However, unchecked or excessive stimulation of the innate immune response can be harmful. It can contribute to the virulence of pathogenic viruses, in part by causing excessive infiltration of the tissues by immune cells, resulting in tissue destruction[5].

Kobasa *et al.*[3] conducted comparative virological and functional-genomic analyses of monkeys infected with either a contemporary influenza virus or the 1918 pandemic virus. Influenza virus strains are generally classified by the two proteins that vary most widely among them — haemagglutinin (H) and neuraminidase (N). Both of the viruses in Kobasa and colleagues' study are of the H1N1 type, although they differ in their epidemiological and pathological properties because of variations in their other proteins.

Analysis of the animals infected with the contemporary virus showed that the virus was present in only a small area of the respiratory tract, causing mild symptoms. Gene expression patterns in the cells lining the bronchus confirmed a robust innate immune response soon after infection. This response included the induction of α/β-interferons, cytokines and chemokines, but the expression of each of these proteins was transient and correlated directly with virus levels in the infected tissues. The authors therefore conclude that infection with contemporary H1N1 influenza virus elicits a transient but appropriate activation of innate immune defences that ultimately facilitates clearance of the virus and recovery.

By comparison, the 1918 virus replicated to high levels and spread rapidly throughout the respiratory tract of infected animals. Lung tissue from these animals showed severe

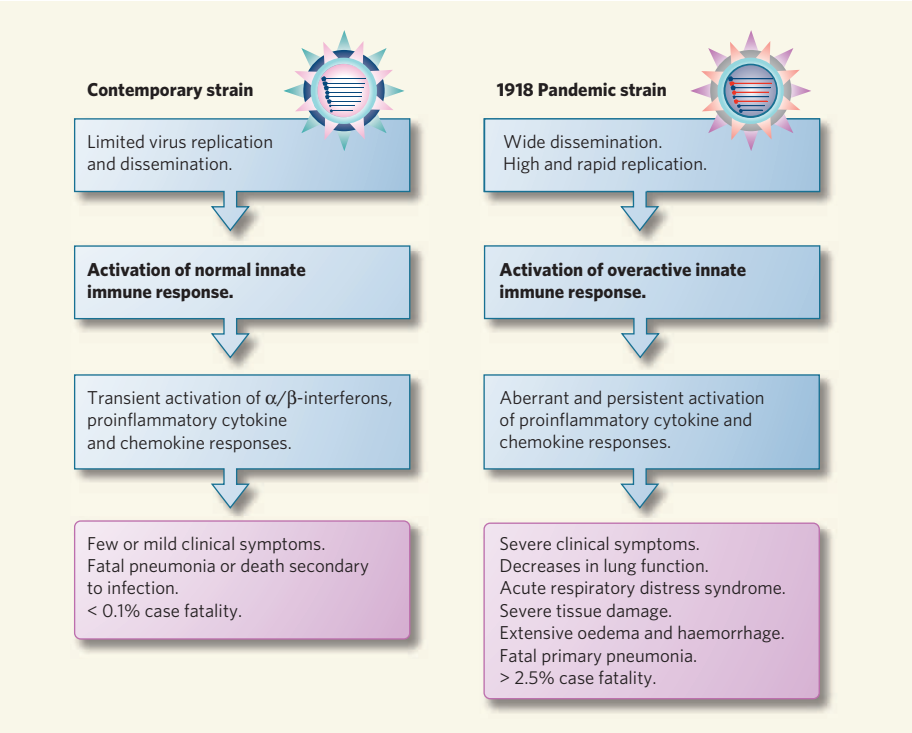


Figure 1 | Different responses to influenza A virus infections. The innate immune response provides protection against virus infection and limits viral disease. Contemporary influenza virus strains induce a transient innate immune response that typically follows a course of self-limiting infection, producing only mild symptoms in patients. Serious disease from contemporary influenza virus is often due to secondary causes. However, the reconstructed 1918 influenza A virus aggressively replicates and disseminates in monkey and mouse models of infection[3,10], and induces an aberrant innate immune response. This could be the cause of the increased severity of symptoms and more deaths in infected people. Blue, observations in animal models; purple, clinical observations in humans.

©2007 Nature Publishing Group

damage, including bleeding and infiltration of immune cells, at times well after the peak of virus replication. A whole-genome analysis of gene expression — the first in monkeys infected with the 1918 virus — showed that the virus triggered aberrantly high and sustained expression of genes encoding many proteins involved in the innate immune response, including proinflammatory cytokines and chemokines. So Kobasa *et al.*[3] propose that an aberrant innate immune response to the 1918 influenza virus could be responsible for the rapid, severe outcome of the infection (Fig. 1). Their data suggest that persistent elevation of inflammatory-response genes could account for the massive inflammation and infiltration of immune cells observed in the respiratory tract of animals infected with the 1918 virus.

Notably, infection with contemporary influenza virus triggered robust expression of genes encoding several α-interferon subtypes and of interferon-stimulated genes with known antiviral properties, and this expression was associated with the comparative mildness of the disease. By contrast, the 1918 virus induced only low and selective expression of genes encoding α-interferon subtypes and differential expression of many interferon-stimulated genes — including one that encodes a protein called RIG-I. Contemporary influenza A virus triggers innate immune defences in part through RIG-I, which regulates the expression of other immune and inflammation genes[6]. But tissue infected with the 1918 virus showed reduced RIG-I expression compared with tissue infected with the contemporary virus. Although these observations may conceptually link RIG-I activity to intracellular pathways that would normally induce the production of interleukins and chemoattractants to clear the virus, the study lacks the biochemical data that are essential for defining such connections.

Kobasa *et al.*[3] demonstrate the power of functional genomics in untangling the complexities of virus–host interactions and viral pathogenesis. Their gene-expression profile analyses suggest that the 1918 virus triggers innate immune signalling processes that possess altered kinetics relative to the contemporary influenza A virus, and/or that the 1918 virus may selectively attenuate the expression of specific innate-response genes. The authors did not, however, analyse gene expression during the important early time course of these events (the first hours and days of infection), so determining the exact mechanisms regulating gene expression will take further study. But it is worth mentioning one possible candidate molecule: the nonstructural-1 (NS1) protein of the influenza A virus can suppress innate immunity by disrupting the induction of α/β-interferon and/or altering the maturation of host-cell RNA[7,8]. Of course, NS1 does not work alone[9], and virulence attributed to pandemic influenza requires many viral components in a unique assemblage of genes[1,2].

The work of Kobasa *et al.* substantiates the findings of Kash *et al.*[10], who showed in mice that the 1918 virus triggered a vigorous innate immune response that was linked to fatalities. Although the mechanisms of tissue destruction were not addressed in either study, the work clearly demonstrates the vital function of early innate immune defences in controlling the virus. It seems that the pandemic 1918 virus had a genetic composition and rapid replication kinetics that may have resulted in an excessively vigorous innate immune and inflammatory response that contributed to severe tissue damage, disease and death.

These conclusions correspond to the striking epidemiological data showing that, unlike contemporary influenza strains, which typically affect the very young and the elderly most severely, the 1918 influenza pandemic was mostly fatal in young adults, who generally possess more robust immune systems[1]. Unveiling the contribution of an aberrant host response to the pathogenesis of the 1918 virus is just the beginning of efforts to understand the disease mechanisms underlying the 1918 pandemic and new virulent strains of influenza virus. The emergence of the H5N1 avian influenza or 'bird flu' virus, and its transfer to the human population, are real and continuing threats[1] that underscore the importance of the current study and of characterizing highly pathogenic forms of flu virus. A better understanding of the origin, transmission and virulence of pandemic influenza viruses, and their interactions with host immune processes, will assist our preparation against future and possibly deadly influenza pandemics. ■

Yueh-Ming Loo and Michael Gale Jr are in the Department of Microbiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9048, USA.
e-mail: michael.gale@utsouthwestern.edu

1. Palese, P. *Nature Med.* **10,** S82–S87 (2004).
2. Tumpey, T. M. *et al. Science* **310,** 77–80 (2005).
3. Kobasa, D. *et al. Nature* **445,** 319–323 (2007).
4. Saito, T. & Gale, M. *Curr. Opin. Immunol.* doi:10.1016/j.coi.2006.11.003 (2006).
5. Wang, T. *et al. Nature Med.* **10,** 1366–1373 (2004).
6. Kato, H. *et al. Nature* **441,** 101–105 (2006).
7. Noah, D. L., Twu, K. Y. & Krug, R. M. *Virology* **307,** 386–395 (2003).
8. Mibayashi, M. *et al. J. Virol.* **81,** 514–524 (2007).
9. Basler, C. F. *et al. Proc. Natl Acad. Sci. USA* **98,** 2746–2751 (2001).
10. Kash, J. C. *et al. Nature* **443,** 578–581 (2006).

SEMICONDUCTOR ELECTRONICS

Trapped fast at the gate

Gerwin Gelinck

The speed record for programming organic transistor memory has been shattered. Work is needed on the stability of the memory storage, but it's a promising step towards some novel technological applications.

The advent of non-volatile flash memory — semiconductor memory that does not lose its data when the power is turned off — revolutionized consumer electronics. It is now used to store the numbers in mobile phones, the pictures taken with digital cameras, and the music tracks in MP3 players. Similar types of memory based on organic semiconductors, rather than traditional silicon-based semiconductors, would make possible entirely new concepts: intelligent food packaging, for instance, that could be used by retailers to control their inventories and to alert consumers when the food is getting close to its 'use-by' date. Writing in *Advanced Materials*, Baeg *et al.*[1] describe an organic thin-film memory transistor that brings such intriguing possibilities a little closer.

Two particular advantages of organic thin-film transistors over solid-state transistors are that they are simpler to make, and can be fabricated on thin, flexible plastic substrates. They could thus form the backbone of low-cost microelectronics ranging from radio-frequency identification tags to flexible, large-area active-matrix displays[2]. Many of these applications require non-volatile (stable) data storage, preferably with memory elements that can be programmed, erased and read electrically. Baeg and colleagues' transistor[1] not only fulfils these requirements, but can also, owing to its similar architecture, be simply integrated into existing technology based on organic transistors.

In order to function as a memory, a device must be observed in two (or more) different states. In flash memories, this is achieved by introducing a second 'floating gate' to the silicon transistor between the normal control gate, which regulates the flow of current through the transistor, and the semiconducting substrate (Fig. 1a). This floating gate is insulated all around by an oxide layer. A high-voltage pulse applied to the control gate places charge on the floating gate, where it becomes trapped. This partially cancels out the electric field coming from the control gate, and so modifies the threshold voltage of the transistor — that is, the voltage required before it lets current flow. Thus, when the transistor is 'read' by placing a specific voltage on the control gate, electrical current will either flow or not flow, depending on the threshold voltage, and so the number of electrons on the floating gate. The resulting

©2007 Nature Publishing Group

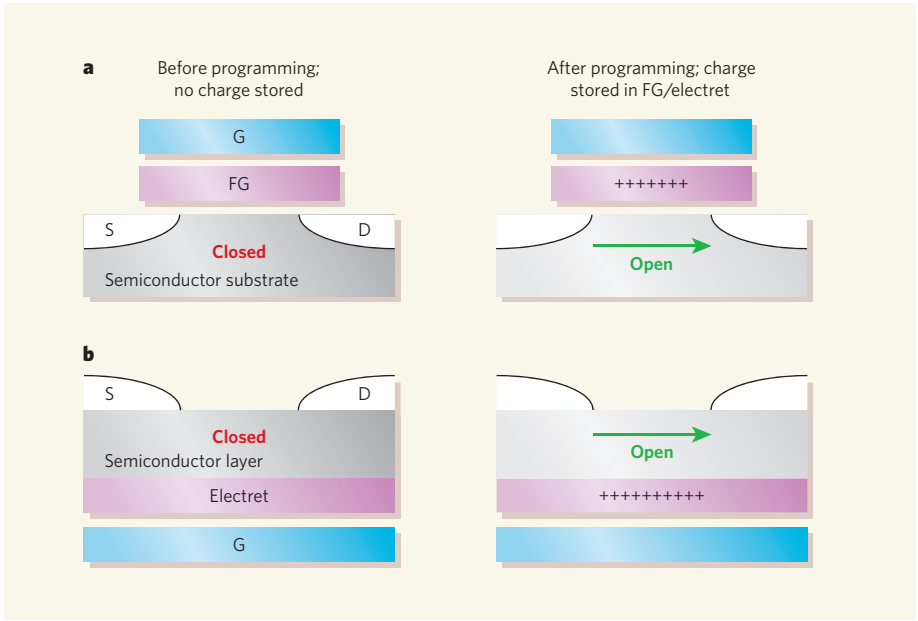


Figure 1 | Methods against memory loss. The basic transistor is a device in which a small voltage applied at the control gate (G) modulates a much larger current flow from source (S) to drain (D) through a semiconductor substrate. **a**, In flash memories, an amount of charge is trapped on a floating gate (FG) that modifies the control voltage required for current to flow from S to D. Whether current flows or not defines a boolean '1' or '0'. The memory of this state persists as long as the charge remains trapped on the floating gate. **b**, In Baeg and colleagues' organic device[1], the same principle is used, but the charge is trapped locally on a thin 'electret' of chargeable polymer, rather than on an isolated floating gate.

current level flowing through the transistor can be used to define boolean '0' and '1' memory states. The charge trapped on the floating gate persists long after all voltages are removed: the information is thus retained after power shutdown, and the transistor functions as a non-volatile memory device.

The non-volatile memory technology developed by Baeg *et al.*[1] takes a slightly different approach (Fig. 1b). Instead of charge being trapped on a floating gate, it is trapped locally on a thin layer of a chargeable polymer. This 'electret' is inserted between the insulating material (silicon dioxide) that makes up the gate and the organic semiconductor (which is pentacene, $C_{22}H_{14}$). When a high-voltage pulse is applied to this electret, it charges up. Applying a reverse voltage pulse either decharges it, restoring the initial state, or charges it to the opposite polarity. The trapped charge imposes an added voltage on the threshold gate voltage, as a floating gate does in a flash memory. This too can be sensed and translated into '1's and '0's, reproducing the stored data, by measuring the current flowing through the transistor.

Memory based on such organic field-effect transistors (OFETs) has been investigated before[3,4], but Baeg *et al.*[1] are the first to report programming speeds of around 1 microsecond — a million times faster than the previous best time of around a second. That marks a decisive step towards making organic memory technology fit for technological purposes.

The new speed record is the result of fast and efficient charge transfer from the organic semiconductor into the polymer electret by means of the electric gate field. The gate field lowers the energy barrier at the interface of the semiconductor and the electret, and so facilitates charge transfer. Once transferred, most of the charge is trapped deeply in the electret. This model explains Baeg and colleagues' most important observations: a critical gate field, caused by the energy barrier, for transfer and trapping of the charges; the lowering of this critical field when the device is illuminated with visible light; a long retention time of the order of hours in the dark; and a decrease of the retention time upon illumination.

These observations also indicate moot points. Can these devices reach data retention times necessary for practical applications — typically years for non-volatile memory? Furthermore, can the device be scaled down to more practical voltages — to 10 V from the 100 V used in the present work — without sacrificing device speed and stability? Such questions remain unanswered, but these encouraging results will without doubt spur intensive investigations into this approach.

The relevance of these results could also go beyond the scope of just memory. Baeg and colleagues[1] aim to produce non-volatile memories that exploit charge trapping and storage, but others are concerned with the converse problem: eliminating charge trapping as much as possible where it gives rise to undesired shifts in threshold voltages that can suppress 'n-channel' (electron) mobility[5] in OFETs, and limit the operational lifetime of 'p-type' (electron-hole-based) logic circuitry[6]. Whatever the intent, all will benefit from a thorough understanding of

50 YEARS AGO

Sir Charles Darwin writes: "The first estimate of Avogadro's number is due to Maxwell himself", and expresses his astonishment that Maxwell "should have published a fact of such tremendous importance in a manner that cannot have drawn much attention to it"... The reason for Maxwell's choice seems to have been that he did not claim to communicate anything fundamentally new, but only to discuss a line of reasoning which Loschmidt had published eight years earlier in the *Proceedings* of the Vienna Academy... It is somewhat surprising that Loschmidt's brilliant achievement has been overlooked so frequently, in spite of Maxwell's full acknowledgement... Avogadro did not even know the order of magnitude of this figure approximately; he died nine years before Loschmidt's paper appeared. **F. A. Paneth**

I must plead guilty to the charge of not having made a very deep search of the older literature in connexion with the evaluation of 'Avogadro's Number'... I am grateful to Prof. Paneth for putting this matter right. **C. G. Darwin**

From *Nature* 19 January 1957.

100 YEARS AGO

The Future in America—a Search after Realities. By H. G. Wells — There has always been in America a wide-spread contempt, not for the law, but for abstract justice, so that even well-minded, influential people do not set themselves to remedy obvious wrong when by doing so they might hurt themselves or their party in the eyes of multitudes of base and busy, greedy and childish, malevolent and ignorant voters. The unfairness of the southerner to the negro is no longer confined to the south, and the crimes of a few negroes exasperate white people so much that they forget the kindly ways of the average man of colour, and thus the negro question is becoming more complex.

From *Nature* 17 January 1907.

50 & 100 YEARS AGO

©2007 Nature Publishing Group

charge-trapping effects at the crucial interface between a semiconductor and its gate. ■

Gerwin Gelinck is at the Holst Centre (Netherlands Organisation for Applied Scientific Research, TNO), High Tech Campus 48, 5656 AE Eindhoven, The Netherlands.
e-mail: gerwin.gelinck@tno.nl

1. Baeg, K.-J. *et al. Adv. Mater.* **18,** 3179–3183 (2006).
2. Dimitrakopoulos, C. D. & Malenfant, P. R. L. *Adv. Mater.* **14,** 99–117 (2002).
3. Katz, H. E. *J. Appl. Phys.* **91,** 1572–1576 (2002).
4. Singh, Th. B. *et al. Appl. Phys. Lett.* **85,** 5409–5411 (2004).
5. Chua, L.-L. *et al. Nature* **434,** 194–199 (2005).
6. Knipp, D., Street, R. A. & Volkel, A. *J. Appl. Phys.* **93,** 347–355 (2003).

CLIMATE CHANGE

Lessons from a distant monsoon

Jonathan T. Overpeck and Julia E. Cole

The burden of global warming falls most heavily on the developing world. A connection forged between the Indian Ocean climate, Asian monsoons and drought in Indonesia makes for an especially bleak outlook for that nation.

As Earth's climate continues to warm, understanding the dimensions of our vulnerability to present and future changes is crucial if we are to plan and adapt. Studies of palaeoclimate have an important role here in helping us to uncover the full range of past climate variability, and so avoid future surprises. On page 299 of this issue, Abram *et al.*[1] present a study of past climate change in Indonesia that expands our view of the pivotal climatological influences in that region to include a geographically distant player: the Asian monsoon.

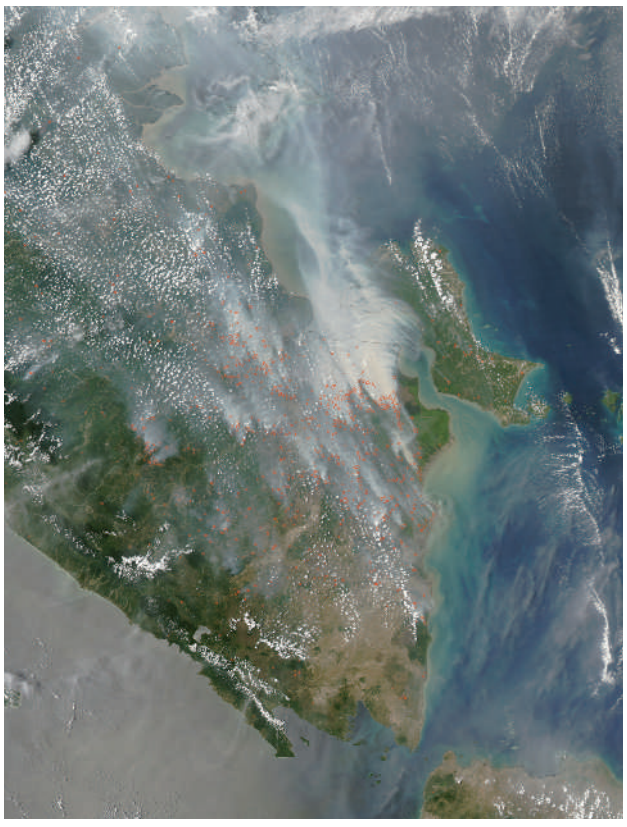
Indonesia's climate is known to vary significantly from year to year as a result of the El Niño/Southern Oscillation (ENSO) system. This system is associated with changes in sea surface temperature and atmospheric pressure across the tropical Pacific. When the central tropical Pacific to the east of Indonesia is warm (an El Niño phase), the normally abundant rainfall in Indonesia moves eastward, leaving much of the island nation in drought. In the west of the country, drought is also brought about by another coupled oscillation in ocean–atmosphere conditions, the 'Indian Ocean Dipole', as a result of cool sea surface temperatures off Sumatra, the most westerly of Indonesia's principal islands.

Abram and colleagues[1] exploit the fact that climate information is preserved in the geochemistry of huge, rapidly growing corals off Sumatra to study past dipole events in the Indian Ocean. Different aspects of coral geochemistry reflect variations in temperature and in the hydrological balance (the difference between levels of precipitation and evaporation). By analysing several geochemical tracers — oxygen isotopic ratios and the ratio of strontium to calcium — in annually banded coral skeletons, the authors can reconstruct month-by-month changes in temperature and drought. Using fossil corals from the mid-Holocene (between around 6,500 and 4,000 years ago), when the Asian monsoon was stronger and ENSO seemingly weaker than today, they demonstrate that the cool ocean temperatures persisted longer — for five months, instead of three — and were accompanied by longer droughts than has been the case in modern times.

Results of climate simulations for 6,000 years ago agree with these observations and suggest a mechanism for the change. First, a stronger Asian monsoon generates anomalies in the easterly winds that would cool the eastern Indian Ocean, predisposing cooler, deep-ocean water in this area to move upward earlier during dipole events. Cooler sea surface temperatures would lead to anomalous downward movement and outflow of air from the region. The resultant weakening of the degree to which moist air converges, together with the atmospheric vertical motion, would result in reduced precipitation, and drought over adjoining land areas.

But what about ENSO? During El Niño conditions, the eastward migration of rainfall and warm ocean temperatures in the tropical Pacific lead to drought in Indonesia. But it has been suggested[2,3] that the mid-Holocene experienced background conditions that may have more closely resembled La Niña conditions. La Niña brings cooler temperatures to the central tropical Pacific, and Indonesia generally receives enhanced rainfall during these periods. So why were droughts more prominent during the mid-Holocene? The implication of Abram and colleagues' work[1] is that the Asian monsoon trumps ENSO and generates prolonged droughts in Indonesia through its influence on the Indian Ocean Dipole. Whether the more frequent droughts associated with interannual variations in ENSO are similarly affected by a stronger monsoon remains unexplored.

The implications of this study for future climate conditions are sobering. If the consensus holds true that the Asian monsoon will intensify with climate warming, Indonesia can expect more frequent and longer droughts in the future through the coupling of the monsoon with sea temperatures in the eastern Indian Ocean. Rural livelihoods and natural resources will thus be at greater risk as drought undercuts regional food supplies and stokes wildfires that also generate exceedingly poor air quality in the region (Fig. 1). Longer droughts will have many additional social and economic consequences, for example on food supply, health and hydropower. Indonesia is also a pivotal



MODIS/NASA/GSFC

Figure 1 | Indonesian summer. Smoke from wildfires on Sumatra, Indonesia, captured on 27 September 2005 by the MODIS (Moderate Resolution Imaging Spectroradiometer) instrument aboard NASA's Aqua satellite. The extent of such fires is likely to increase if climate warming causes the Asian monsoon to intensify.

©2007 Nature Publishing Group

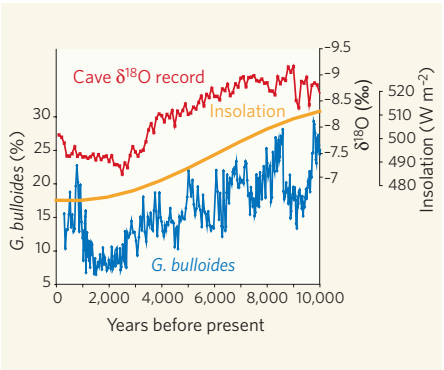


Figure 2 | Abrupt changes in monsoon climate. The weakening of the monsoon during the Holocene period is shown here by the relative abundance of the planktonic foraminiferan *Globigerina bulloides* in the Arabian Sea[11] (blue line), which reflects the intensity of the upwelling of sea water and the strength of monsoon winds. Monsoon weakening is also indicated by records of the oxygen isotopic ratio $\delta^{18}O$ (red line, average of three records) from Dongge Cave in southern China, reflecting the relative amount of summer rainfall[12–14]. At both sites, monsoon weakening is forced primarily by slow changes in summer sunlight[15] ('insolation', yellow; calculated for June at 30° N latitude), but apparently can occur abruptly as a series of steps. Abrupt changes in the Asian monsoon could impinge directly on Indonesia's climate, as Abram *et al.*[1] show.

biodiversity hotspot[4], and drought intensification could bring significant challenges to conservation management.

The troubling consequences do not end there. The palaeoclimate record seems to indicate that the Asian monsoon does not always respond linearly to climate change. The main driver of large-scale monsoon change over the past 10,000 years has been a slow decrease in summer-time solar radiation (insolation) owing to changes in Earth's orbit. Even so, monsoon records show abundant evidence of abrupt, stepwise changes on timescales of a century and shorter (Fig. 2). The mechanisms behind this behaviour are poorly understood, and research is continuing in an attempt to document the degree to which abrupt monsoon changes are spatially and temporally coherent[5].

Nevertheless, palaeoclimate observations caution us to expect similarly abrupt changes as Earth's climate warms. As predictive models do not yet simulate the past abrupt changes in the Asian monsoon, they cannot be trusted to project such abrupt changes realistically into the future. Future monsoon behaviour will also depend on how such natural variations interact with the combined anthropogenic effect from aerosols (thought to weaken monsoon intensity) and greenhouse gases (thought to strengthen it). For example, reductions in air pollution in Asia could, by reducing the damping effects of aerosols on the monsoon, result in sudden monsoon strengthening[1,6].

Indonesia faces the possibility of climate surprises on many fronts. As a nation of islands, it is particularly vulnerable to the rises in sea level of a few metres that are possible if climate warming reaches the 'tipping point' for Earth's ice sheets[7]. Evidence is mounting that a warmer world will experience tropical storms of greater intensity[8,9]. The future behaviour of ENSO is uncertain[10], but also critical for regional climate projections, nowhere more so than in Indonesia. As the regional nuances of future environmental change come into better focus, the dimensions of Indonesia's, and other nations', vulnerability will become slowly clearer. ■

Jonathan T. Overpeck and Julia E. Cole are in the Departments of Geosciences and Atmospheric Sciences, University of Arizona, Tucson, Arizona 85721, USA.
e-mail: jto@u.arizona.edu

1. Abram, N. J. *et al. Nature* **445,** 299–302 (2007).
2. Clement, A. C., Seager, R. & Cane, M. A. *Paleoceanography* **15,** 731–737 (2000).
3. Cole, J. E. *Science* **291,** 1496–1497 (2001).
4. Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. *Nature* **403,** 853–858 (2000).
5. Morrill, C., Overpeck, J. T. & Cole, J. E. *Holocene* **13,** 465–476 (2003).
6. Meehl, G. A. *et al. Geophys. Res. Lett.* **23,** 3755–3758 (1996).
7. Overpeck, J. T. *et al. Science* **311,** 1747–1750 (2006).
8. Webster, P. J., Holland, G. J., Curry, J. A. & Chang, H. R. *Science* **309,** 1844–1846 (2005).
9. Emanuel, K. *Nature* **436,** 686–688 (2005).
10. Collins, M. *et al. Clim. Dynam.* **24,** 89–104 (2005).
11. Gupta, A. K. & Anderson, D. M. *Geophys. Res. Lett.* **32,** doi:10.1029/2005GL022685 (2005).
12. Wang, Y. J. *et al. Science* **308,** 854–857 (2005).
13. Dykoski, C. A. *et al. Earth Planet. Sci. Lett.* **233,** 71–86 (2005).
14. Yuan, D. X. *et al. Science* **304,** 575–578 (2004).
15. Berger, A. & Loutre, M. F. *Quat. Sci. Rev.* **10,** 297–317 (1991).

POLYMER CHEMISTRY

Sacrificial synthesis

Sébastien Perrier and Xiaosong Wang

The size and uniformity of polymer molecules makes it difficult to modify them at just one selected site. But a single chemical group can be attached at the end of a polymer if part of the starting material is forfeited.

Chemists are remarkably proficient at directing the synthesis of small molecules, but fine-tuning the structures of large molecules, such as polymers, is far more taxing. Despite many years of research, the field of macromolecular engineering — the preparation of large molecules with strict control over their size and chemical groups — has many mountainous challenges yet to overcome. But Kilbinger and colleagues[1] have just conquered a particularly troublesome peak. Reporting in *Angewandte Chemie*, they describe an innovative approach for preparing polymer chains with just one chemical group attached at the end, with excellent control over the chain length. Such molecules are much sought-after for potential applications in nanotechnology.

Nature provides endless examples of precisely engineered macromolecules — proteins, for instance, which contain amino-acid side-chains that are accurately positioned, often in a way that determines the proteins' roles. Synthetic chemists have tried to recreate nature's exceptional control over macromolecules, and in so doing they have designed several catalytic reactions that occur only at specific chemical groups.

One such reaction is known as ring-opening metathesis polymerization (ROMP). In this process, an organometallic catalyst breaks the carbon–carbon double bond in a cyclic molecule, to form a ring-opened structure that has a carbon–carbon double bond at one end and a metal–carbon double bond (known as a metal carbene) at the other. The metal carbene can react in the same way as the catalyst, so that a long chain propagates as more ring-opened molecules are added, much as a daisy chain is formed from individual flowers. Eventually, very long polymer chains are formed that contain carbon–carbon double bonds in their 'backbones' and that have a metal carbene at one end. The metathesis reaction that underpins this process is astonishingly versatile, and its pioneers won the Nobel Prize in Chemistry for their work in this area[2–4].

As ROMP leads to a polymer with an active metal carbene at one end, many researchers have attempted to convert this reactive terminus into other useful chemical groups. The product of this reaction is highly desirable: a polymer with just one reactive group attached. Such structures are of interest in molecular engineering because they are potential building blocks for nanocomposite materials, or for biological–synthetic macromolecules (for example, hybrids of proteins and polymers). However, all the attempted reactions with metal carbenes at polymer terminals were plagued by low levels of conversion into the desired product, mainly due to side reactions. For instance, the metal carbene can easily react with double bonds in the polymer chains, causing unwanted chain transfers[5].

Historically, the method of choice for generating mono-derivatized polymers has been anion-initiated polymerization (rather than the metal-carbene-initiated ROMP process). In this process, the active chain-ends of the polymers can be reacted with chemical

©2007 Nature Publishing Group

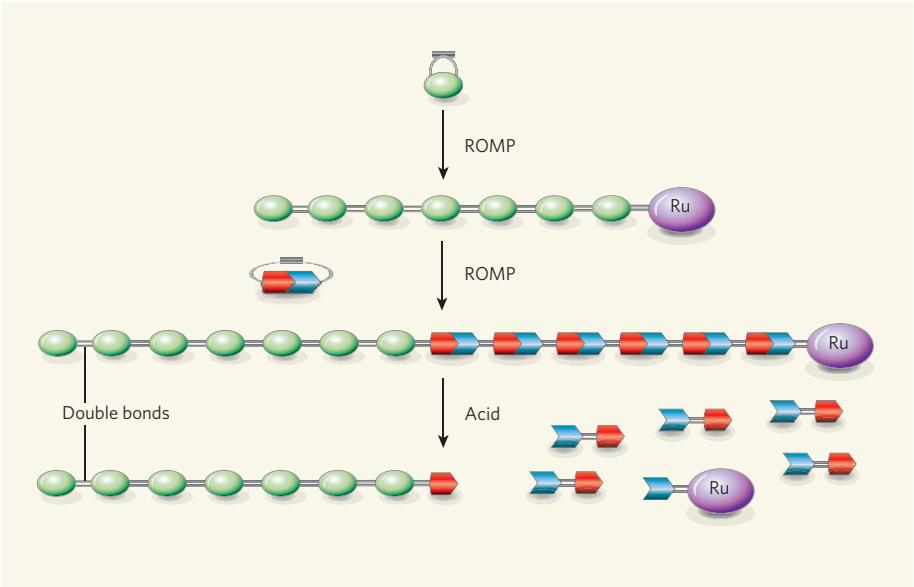


Figure 1 | Sacrificial polymers. Kilbinger and colleagues[1] have devised a strategy for preparing polymers with a single chemical group attached at one end. A polymer chain is made in a ring-opening metathesis polymerization (ROMP) reaction, which leaves a reactive metal complex — in this case, a ruthenium (Ru) complex — attached to the product by a double bond. This product undergoes another ROMP reaction with a different monomer, to make a chain made of two distinct polymer sections. Upon acid treatment, the second section decomposes, leaving the first polymer chain intact with a single chemical group appended.

terminating agents, so introducing the required single group. Alternatively, the anionic initiator may incorporate a masked chemical group that can be exposed once the polymerization is complete. But these polymerization reactions require stringent reaction conditions — such as very low temperatures and extremely pure chemicals — that make the synthesis cumbersome. Furthermore, anions will react with most chemical groups, so the choice of terminating groups that will survive the reaction conditions is limited; multi-step reactions are usually required to obtain desirable end-groups[6].

But now, Kilbinger and colleagues[1] have reported a fascinating strategy to circumvent these problems. Their idea is to generate mono-derivatized polymers from ROMP reactions by using a 'block copolymer' as a precursor. The molecules of a block copolymer consist of two or more sections (blocks), where each section is a different kind of polymer. The authors proposed using a copolymer consisting of two blocks, the first of which is the target polymer, and the second of which is a sacrificial chain. The material for the second segment must be able to decompose under specific conditions, leaving a single chemical group behind on the terminal of the first block (Fig. 1). This approach offers two advantages over the existing methods. First, the side reactions described previously can be avoided. Second, as block copolymers can easily be purified, it is straightforward to obtain a very high proportion of polymers with a single chemical group at their chain ends.

To demonstrate their approach, the authors used dioxepine molecules — which contain chemical groups known as acetals — as ROMP monomers to construct their sacrificial polymer. The resulting block incorporates the acetal groups into its backbone. Kilbinger and colleagues showed that these backbone acetal groups decompose under acidic conditions, destroying the sacrifical block and leaving a residual hydroxyl group (OH) at the end of the target polymer. They went on to demonstrate that this hydroxyl group is a useful chemical starting point by converting it directly into other chemical groups using well-established reactions. This elegant approach provides a versatile method for the synthesis of mono-substituted polymers, with exquisite control over the size of the molecules.

So what impact will Kilbinger and colleagues' method make? Recent advances in 'living radical polymerization' (LRP) have essentially solved the problems associated with the anionic approach. The LRP process is based on free-radical polymerization, which is compatible with many chemical groups and can be performed under much milder conditions than anionic polymerization. Using initiators that incorporate specific chemical groups, polymers decorated with many chemical motifs can be produced in a straightforward manner[7]. The drawback is that LRP is only possible with a limited number of monomers. Given such limitations, and the many industrial applications of double-bond-containing polymers (such as those produced by ROMP), Kilbinger and colleagues' development will undoubtedly expand the toolbox for macromolecular engineering. ■

Sébastien Perrier and Xiaosong Wang are in the Department of Colour and Polymer Chemistry, University of Leeds, Leeds LS2 9JT, UK.
e-mails: s.perrier@leeds.ac.uk; x.wang@leeds.ac.uk.

1. Hilf, S., Berger-Nicoletti, E., Grubbs, R. H. & Kilbinger, A. F. M. *Angew. Chem. Int. Edn* **45,** 8045–8048 (2006).
2. Chauvin, Y. *Angew. Chem. Int. Edn* **45,** 3740–3747 (2006).
3. Schrock, R. R. *Angew. Chem. Int. Edn* **45,** 3748–3759 (2006).
4. Grubbs, R. H. *Angew. Chem. Int. Edn* **45,** 3760–3765 (2006).
5. Grubbs, R. H. (ed.) *Handbook of Metathesis* Vol. 3 (Wiley-VCH, Weinheim, 2003).
6. Szwarc, M. *Nature* **178,** 1168–1169 (1956).
7. Matyjaszewski, K. (ed.) *Controlled/Living Radical Polymerization: from Synthesis to Materials* (American Chemical Society, Washington DC, 2006).

CELL BIOLOGY

A switch for S phase

Michael Botchan

DNA replication is a necessary prelude to the division of a eukaryotic cell. Initiation of this process requires a complex script, involving many proteins: details of one of the main acts now emerge.

In 1992, a report by Li and Alberts[1] appeared in these pages to comment on a seminal discovery[2] — the identification of the protein factors from the budding yeast *Saccharomyces cerevisiae* that recognize the spots on DNA for initiating the DNA-replication phase of the cell cycle. Finding this 'origin-recognition complex' was one task, however; identifying the participants in the actual switch that starts DNA synthesis is another. This latter achievement is now described in papers by Zegerman and Diffley[3], and by Tanaka *et al.*[4], on pages 281 and 328 of this issue. At the heart of this intricate process is the phosphorylation — and thus activation — of two proteins known as Sld3 and Sld2, and the involvement of a third named Dpb11.

The pathway to DNA replication begins early in the cell cycle, with events occurring at the end of nuclear division and in the ensuing G1 phase. DNA synthesis — during which a complete single copy of the entire genome is made — then commences in S phase. Transitions between all cell-cycle phases are controlled by the activation and deactivation of a series of cyclin-dependent kinases (CDKs),

©2007 Nature Publishing Group

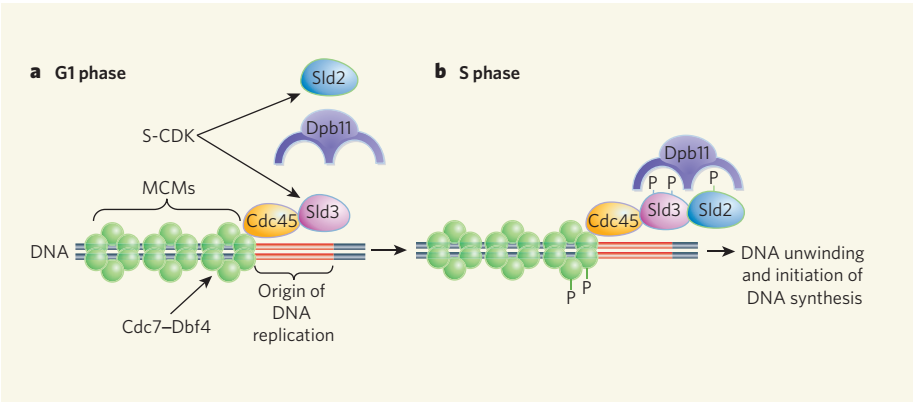


Figure 1 | DNA replication and the switch from G1 to S phase. **a**, During G1 phase, a pre-replication complex is assembled close to the origin of DNA replication. The MCM proteins are loaded onto the site by the action of the origin-recognition complex and Cdc6 (not shown, but positioned over the origin of DNA replication). Sld3 and Cdc45 associate in an undefined manner with the DNA–protein complex. **b**, DNA synthesis starts rapidly after two kinases shown in **a** — S-CDK and Cdc7–Dbf4 — phosphorylate (P) the other factors as indicated. Both of these kinases interact with the origin-recognition complex, which targets them to the origin site. Modifications bring Sld3 and Sld2 together with Dpb11 (through Dpb11's BRCT repeats). The assembly of these proteins, and perhaps others, at an origin of DNA replication starts the process of unwinding the DNA, and then DNA synthesis, by as-yet unknown mechanisms. The upshot of the new work[3,4] has been to identify the docking of Sld3 and Sld2 to Dpb11 as the essential events in the switch.

which control the phosphorylation of other proteins. Thus, after the origin-recognition complex had been identified, finding the actual targets for S-CDK, the CDK known to promote the switch from G1 to S phase, became a major objective.

Understanding the switch required investigations of the complex pathway that brings the almost completed replication machinery to the correct position on the DNA. What emerged first was an understanding of the assembly of a pre-replication complex during G1 (Fig. 1). This requires many of the factors that interact with the origin-recognition complex and low levels of CDK. The step of unwinding the DNA strands, which is carried out by a helicase enzyme, must be one of the earliest steps in DNA synthesis, and an inactive form of such an enzyme arrives at the pre-replication complex early in the pathway. The origin-recognition complex and another player, Cdc6, form a machine that recruits other proteins (minichromosome maintenance proteins, or MCMs) to the origin of DNA replication during G1 (ref. 5); the MCMs eventually provide essential helicase activity during S phase. Other studies[6] have revealed how components of the pre-replication complex are destroyed by the S-phase-promoting factors, so allowing for the once-and-only-once initiation of DNA synthesis. But the actual switch to start synthesis remained elusive.

Zegerman and Diffley[3] and Tanaka *et al.*[4] now reveal the components of that switch — that is, the targets of the S-CDKs — in *S. cerevisiae*. They also disclose surprising details about what must be accomplished during G1 for DNA synthesis to begin. The switch probably occurs with the activation of the DNA unwinding process that creates the substrates for polymerase, the enzyme that does the actual work of synthesis, and involves Dpb11, Sld3 and Sld2 — or 11–3–2 for short.

Previous work by Araki and colleagues had shown that Sld2 is an essential target for S-CDK, and that there are genetic and biochemical interactions between Sld2 and Dpb11. Crucially, the phosphorylated form of Sld2 binds tightly to particular parts of Dpb11 known as BRCT repeats. When replacing a normal copy of the gene, expression of an engineered *SLD2* gene encoding a 'phosphomimetic' residue could maintain function, but cells still required passage through G1 and functioning S-CDK. This finding implied, but did not prove, that there are other targets for S-CDK. Sld3 contains 12 sites for phosphorylation by CDKs, and genetic interactions between *DPB11* and *SLD3* brought Sld3 into the experimental spotlight. The new experiments of Zegerman and Diffley[3] and Tanaka *et al.*[4] show that the phosphorylation of Sld3 on two amino-acid residues, and of Sld2 on a single residue, are all that are required for the switch to operate. Importantly, the two phosphorylated proteins, Sld2 and Sld3, must bind to Dpb11 (Fig. 1), and this docking presumably occurs at sites marked as origins of DNA replication by the origin-recognition complex.

Armed with this knowledge, the two groups[3,4] were able to generate S-CDK 'bypass' conditions, in which there was no need for S-CDK to prompt S phase. This allowed them to tease apart the machinery determining entry into S phase. Their approach involved α-factor, a yeast hormone that arrests cells in G1. Together with genetic manipulations, this factor can be used to investigate the consequences for a strain of yeast that can enter S phase without S-CDK activity.

Under conditions in which such activity is not required, we might have expected premature DNA replication and complete bypass of the G1 phase. We learn from Zegerman and Diffley[3] that α-factor-arrested cells with the S-CDK bypass system show only weak DNA replication. G1 is a busy time for cells and it turns out that another essential kinase, Cdc7, has a regulatory subunit, Dbf4, that is degraded early on and only reaches critical levels after the G1-specific CDKs have inactivated the protein-degradation machinery. Among the main targets for the Cdc7–Dbf4 kinase in DNA replication are the MCM proteins[7] (Fig. 1). Zegerman and Diffley report that, with the bypass system in place and adequate expression of Dbf4, DNA replication can be strong in α-factor G1-arrested cells and, as would be expected, is lethal. So, regulation of both Cdc7 and S-CDK activity is required to prevent premature DNA replication in G1 phase, providing the sequence of events shown in Figure 1. Thus we have a bare skeleton of what must be accomplished during G1 to provide the necessary and sufficient conditions for passage into the next phase of the cell cycle — at least for DNA replication.

What is now needed is a mechanistic insight into what the 11–3–2 complex actually does

Box 1 | The 11–3–2 complex and molecular matchmaking

Clues to what the 11–3–2 (Dpb11–Sld3–Sld2) complex does to initiate DNA replication come from some of Tanaka and colleagues' experiments[4]. They looked for mutants that would be lethal if both the bypass version of Sld2 and high levels of Dbf4 were expressed. The screen uncovered a form of the *CDC45* gene, known as *JET1*, which encodes another protein known to be essential for DNA replication. Expression of *JET1* with the *SLD2*-bypass gene also uncouples the initiation of DNA replication from the G1 phase. Cdc45 is known to sit with the origin-recognition complex, Sld3, Cdc6 and the minichromosome maintenance proteins (MCMs) at the origin of replication (Fig. 1). But the only proteins from this group that travel from this site with the DNA polymerase are the MCMs (with the associated helicase) and Cdc45.

Perhaps the essential step in activating the helicase is the binding of other factors to the MCMs. A guess as to the identity of those factors would be the 'GINS' complex, which is composed of four proteins each essential for DNA replication. The 11–3–2 complex might be a molecular matchmaker that completes the assembly of the entire replication machine that follows the DNA polymerase, and JET1 might stabilize the association of the GINS complex with this apparatus. A complex[9] containing MCMs, Cdc45 and the GINS complex shows helicase activity *in vitro*, and such factors do travel with the DNA polymerase[10,11]. **M.B.**

©2007 Nature Publishing Group

to initiate DNA replication; whatever that function is, it is transient and not required beyond the step that instigates DNA synthesis[8]. Some suggestions are detailed in Box 1. This topic will now be the subject of considerable attention.

Meanwhile, I can echo a comment made 15 years ago. In their News & Views article[1], Li and Alberts equated the successful identification of the origin-recognition complex to the finding of the holy grail. For cell and molecular biologists, that sentiment can be echoed to express the significance of the discoveries made by Zegerman and Diffley[3], and Tanaka and colleagues[4]. ■

Michael Botchan is in the Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, California 94720-3200, USA.
e-mail: mbotchan@berkeley.edu

1. Li, J. J. & Alberts, B. *Nature* **357,** 114–115 (1992).
2. Bell, S. P. & Stillman, B. *Nature* **357,** 128–134 (1992).
3. Zegerman, P. & Diffley, J. F. X. *Nature* **445,** 281–285 (2007).
4. Tanaka, S. *et al. Nature* **445,** 328–332 (2007).
5. Randell, J. C. W., Bowers, J. L., Rodriguez, H. K. & Bell, S. P. *Mol. Cell* **21,** 29–39 (2006).
6. Blow, J. J. & Dutta, A. *Nature Rev. Mol. Cell Biol.* **6,** 476–486 (2005).
7. Sheu, Y. J. & Stillman, B. *Mol. Cell* **24,** 101–113 (2006).
8. Kanemaki, M. & Labib, K. *EMBO J.* **25,** 1753–1763 (2006).
9. Moyer, S. E., Lewis, P. W. & Botchan, M. R. *Proc. Natl Acad. Sci. USA* **103,** 10236–10241 (2006).
10. Pacek, M. *et al. Mol. Cell* **21,** 581–587 (2006).
11. Gambus, A. *et al. Nature Cell Biol.* **8,** 358–366 (2006).

COSMOLOGY

The Universe's skeleton sketched

Eric V. Linder

The deepest and clearest maps yet of the Universe's skeleton of dark-matter structure present a picture broadly in concord with favoured models — although puzzling discrepancies remain.

Astronomers use every resource at their disposal to construct an image of the Universe: from the microwave radiation of the cooling Big Bang, past the visible wavelengths of light seen by the casual viewer of the night sky, to the γ-rays of the most powerful explosions of collapsing stars and ravenous black holes. They don't even have to see an object directly to detect its presence. In a manner akin to the Polynesian seafarers who sense islands out of their sight through the deflected direction of ocean waves, cosmologists can map a concentration of the Universe's unseen mass through the gravitational deflection of light coming from sources behind it. On page 286 of this issue[1], Massey *et al.* publish the latest and most detailed atlas of these 'dark-matter' congregations*.

Dark matter is thought to comprise as-yet undiscovered elementary particles, and is a central prediction of the current 'concordance' model of the Universe's make-up. Here, it acts as a skeleton around which bright matter — galaxies and clusters of galaxies — assembles. There is abundant indirect evidence for structures made of dark matter. The gravitational deflection, or lensing, technique has already been used to make maps of dark matter around individual galaxies and clusters, most dramatically by Clowe *et al.*[2]. Low-resolution reconstructions of the skeleton of a larger chunk of the nearby Universe have also been carried out using ground-based telescopes.

Massey *et al.*[1] present an analysis of gravitational-lensing data from the Cosmic Evolution Survey, COSMOS, that were acquired using the Hubble Space Telescope. These data cover an area of sky about eight times the size of the full Moon, and represent the first wide-sky, space-based survey of dark matter. This is a great step forward, as observations from space avoid the distortions and time variations that Earth's atmosphere imposes on the astronomical signal. (The effect of these perturbations is rather as though our seafarer friend had to contend with a nearby canoe's paddlers roiling the ocean when studying his waves.) The new maps consequently have a much higher resolution than the best ground-based observations, with four times the density of sources than in previous studies. They also go deeper into the Universe, looking back to a time when the Universe was about half the age it is now, at a redshift $z = 1$. (The redshift is a measure of the expansion of the Universe; the higher the redshift of a cosmological object, the smaller and younger the Universe was when the object sent out its signal.)

To assess the third dimension of their map, the depth of the field of view, astronomers cannot move their canoe and use triangulation. They rely instead on mathematics to derive the distance of the gravitational lens by means of its focal length — which is known from Einstein's general theory of relativity — and the distance to the background source galaxy whose light is probing the dark matter. Massey *et al.* combine the COSMOS observations with follow-up observations from ground-based telescopes in 15 wavelength bands, and so are able to estimate redshifts of the galaxy sources all the way back to $z = 3$, when the Universe was just a sixth of its present age.

With this information, the pattern of dark-matter concentrations can be mapped crudely in three dimensions by slicing the data into three redshift shells. Although this is just a first step in 'cosmic tomography', the advantage of using a space telescope is that the source density in each slice of this three-dimensional map is nearly as high as for the 'unsliced' maps — two-dimensional projections — expected from the next generation of ground-based surveys.

Massey and colleagues' three-dimensional map can be studied on its own merits to provide information about the pattern and evolution of dark-matter structures, or can be compared with observations of the bright matter through visible light and X-rays. The dark-matter distribution shows modest evidence for a connected, filamentary distribution — exactly the kind of skeleton predicted by concordance cosmology for stars and galaxies to assemble around.

A note of caution, however: it is still early days, and spurious noise affects the weaker features in the map (at the level of around 2.9 standard deviations). There are also striking examples of both beautiful agreement and puzzling disagreement between the dark and bright maps. Areas bright in X-rays, indicating the highest concentrations of gas heated by falling into the strongest gravitational fields, almost always overlap with the dark-matter concentrations — but not absolutely always. Conversely, dark-matter concentrations sometimes seem to have no corresponding bright matter. For the dark-matter skeleton of mass in the Universe, flesh sometimes occurs without supporting bones, and bones without surrounding flesh. That is perfectly possible. However, these data, together with forthcoming ground-based data from the lower-resolution, but wider-area Canada–France–Hawaii Telescope Legacy Survey[3], will need to be carefully compared with computational simulations of the Universe to analyse consistencies and discrepancies.

Cosmologists are getting a foretaste of a new ability to map the dark Universe, despite the fact that the gravitational-lens mapping technique is still showing some growing pains: even the Hubble Space Telescope observations required significant corrections, because of radiation damage to the detectors. Radiation-resistant detectors have already been developed for a proposed space mission, the Supernova/Acceleration Probe (SNAP)[4], that would extend COSMOS-type maps of the dark-matter skeleton to an area 2,500 times larger than that covered by Massey *et al.*[1]. These maps would also be both deeper and more finely resolved. But for now, the COSMOS data give astronomers an exciting new view of the dark Universe. ■

Eric V. Linder is at the University of California and Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.
e-mail: evlinder@lbl.gov

1. Massey, R. *et al. Nature* **445,** 286–290 (2007).
2. Clowe, D. *et al. Astrophys. J. Lett.* **648,** 109–113 (2006).
3. www.cfht.hawaii.edu/Science/CFHLS
4. http://snap.lbl.gov

*This article and the paper concerned[1] were published online on 7 January 2007.

©2007 Nature Publishing Group

ASTROPHYSICS

Quark matter in compact stars?

Arising from: F. Özel *Nature* 441, 1115–1117 (2006)

In a theoretical interpretation of observational data from the neutron star EXO 0748–676, Özel concludes that quark matter probably does not exist in the centre of neutron stars[1]. However, this conclusion is based on a limited set of possible equations of state for quark matter. Here we compare Özel's observational limits with predictions based on a more comprehensive set of proposed quark-matter equations of state from the literature, and conclude that the presence of quark matter in EXO 0748–676 is not ruled out.

Özel's stated lower limits on the mass and radius are $M \geq 2.1 \pm 0.28\,M_{\odot}$ and $R \geq 13.8 \pm 1.8$ km. She correctly points out that these values exclude a soft equation of state, but then infers that there is no quark matter in this compact star. However, this conclusion is not justified because quark matter can be as stiff as nuclear matter, because effects from strong interactions (quantum chromodynamics, QCD) can harden the equations of state substantially. The corresponding hybrid or quark stars can indeed reach a mass of $2M_{\odot}$, as demonstrated in calculations using the MIT bag model[2], perturbative corrections to QCD[3], and the Nambu–Jona–Lasinio model[4]. The mass–radius relations for compact stars using various quark-matter and nuclear-matter equations of state, together with the lower limits derived by Özel, are shown in Fig. 1.

In addition to the mass and radius, there are potential constraints on (or signatures of) the presence of quark matter from observations of the cooling, spin-down and precession of neutron stars, and from transient phenomena such as glitches, magnetar flares and superbursts.

Cooling observations of firmly identified neutron stars are mostly consistent with a 'minimal' model of nuclear-matter cooling. However, there is evidence of faster cooling in limits obtained from supernova remnants, and the presence of exotic forms of matter is not ruled out[5]. A detailed analysis of cooling data, including information from elliptical flow in heavy-ion collisions, was unable to find any purely nuclear equation of state that was compatible with all the data[6]. Models involving some quark matter in the cores of neutron stars were more successful[7,8].

Measurements of the spin-down rate of neutron stars can be used to constrain the shear and bulk viscosity of the interior, because sufficiently low viscosity would lead to very fast spin-down by gravitational radiation from unstable *r*-modes. Preliminary calculations rule out a strange star made of 'colour–flavour-locked' (CFL) matter[9], but hybrid stars are not ruled out. More controversially, it has also been argued that the measured precession of some

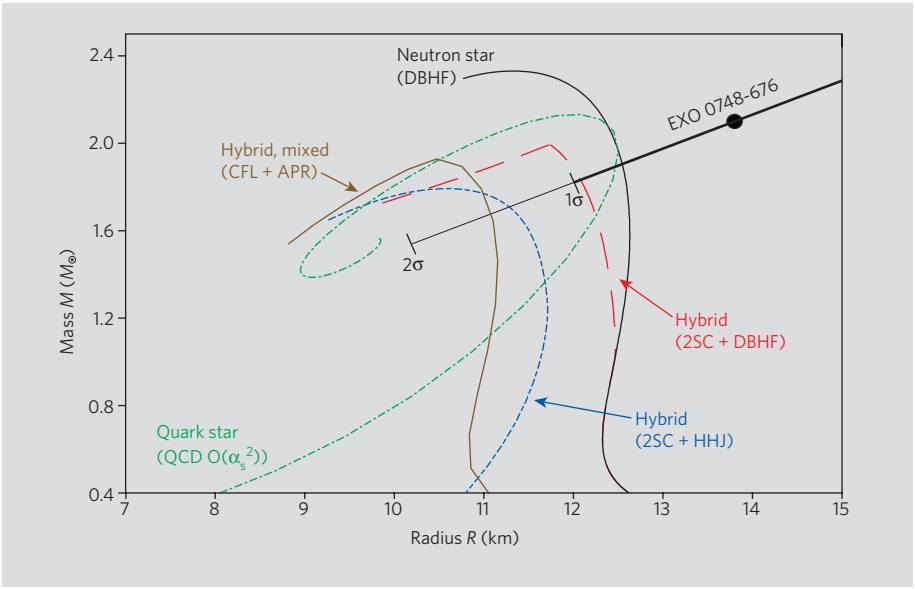


Figure 1 | Lower limit of mass, *M*, and radius, *R*, from Özel's analysis[1] of EXO 0748–676 (1σ and 2σ error bars are shown), and the calculated *M*–*R* curve for various quark- and nuclear-matter equations of state. These include pure nuclear matter described by the DBHF (relativistic Dirac–Brueckner–Hartree–Fock) equation of state[6] (black line); a hybrid star with a core of 2SC quark matter (a two-flavour colour-superconducting phase, in which only the up and down quarks form Cooper pairs) and a mantle of DBHF nuclear matter[8] (red line); a hybrid star with a core of 2SC quark matter and mantle of HHJ nuclear matter (APR with high-density causality corrections)[7] (blue line); a hybrid star whose core is a mixed phase of APR nuclear matter (based on the Argonne v_{18} two-nucleon interaction with variational chain summation) and CFL quark matter (the 'colour–flavour-locked' colour-superconducting phase in which all three colours and flavours undergo Cooper pairing)[2] (brown line); and a pure quark-matter star using an equation of state with order α_s^2 QCD corrections[3] (green line). The presence of quark matter is therefore not excluded by the EXO 0748–676 results.

stars is inconsistent with the standard understanding of nuclear matter[10].

Glitches (temporary speeding-up in the rotation of a neutron star that is gradually spinning down) are only partially understood, but are believed to provide evidence for a substantial crust overlapping with a superfluid region inside the star[11]. This does not exclude the presence of a quark-matter core, and may not even exclude strange stars, as there are superfluid and crystalline phases of quark matter[12,13].

Observations of quasi-periodic oscillations in soft gamma repeaters have been used to obtain the frequencies of toroidal shear modes of their crusts. The results are not consistent with these objects being purely strange stars, but they put no limits on the presence of a quark-matter core inside them[14].

Observations of superbursts in low-mass X-ray binaries yield ignition depths much smaller than those predicted for standard neutron stars, and are more compatible with these objects being hybrid stars with a relatively thin baryonic crust[15].

We conclude that Özel's analysis, assuming that it is indeed correct, can be used to put constraints on the parameters of the quark-matter equations of state, but that neither Özel's analysis nor the other available observational data have yet ruled out the presence of deconfined quarks in compact stars.

M. Alford*, D. Blaschke†‡, A. Drago§||, T. Klähn‡¶, G. Pagliara§||, J. Schaffner-Bielich#
*Department of Physics, Washington University, St Louis, Missouri 63130, USA
e-mail:alford@wuphys.wustl.edu
†Instytut Fizyki Teoretycznej, Uniwersytet Wroclawski, 50-204 Wroclaw, Poland
‡Institut für Physik, Universität Rostock, 18051 Rostock, Germany
§Dipartimento di Fisica, Università di Ferrara, 44100 Ferrara, Italy
||INFN, Sezione di Ferrara, 44100 Ferrara, Italy
¶GSI, 64291 Darmstadt, Germany
#Institut für Theoretische Physik, Goethe Universität, 60438 Frankfurt am Main, Germany

1. Özel, F. *Nature* **441,** 1115-1117 (2006).
2. Alford, M., Braby, M., Paris, M. & Reddy, S. *Astrophys. J.* **629,** 969-978 (2005).
3. Fraga, E. S., Pisarski, R. D. & Schaffner-Bielich, J. *Phys. Rev. D* **63,** 121702 (2001).
4. Rüster, S. B. & Rischke, D. H. *Phys. Rev. D* **69,** 045011 (2004).
5. Page, D., Geppert, U. & Weber, F. *Nucl. Phys. A* **777,** 497-530 (2006).

© 2007 Nature Publishing Group

6. Klähn, T. *et. al. Phys. Rev. C* **74,** 035802 (2006).
7. Grigorian, H., Blaschke, D. & Voskresensky, D. *Phys. Rev. C* **71,** 045801 (2005).
8. Klahn, T. *et al. arXiv:nucl-th/*0609067 (2006).
9. Madsen, J. *Phys. Rev. Lett.* **85,** 10–13 (2000).
10. Link, B. *Phys. Rev. Lett.* **91,** 101101 (2003).
11. Alpar, M. A. *Phys. Rev. Lett.* **58,** 2152 (1987).
12. Alford, M., Rajagopal, K. & Wilczek, F. *Nucl. Phys. B* **537,** 443–458 (1999).
13. Alford, M., Bowers, J. A. & Rajagopal, K. *Phys. Rev. D* **63,** 074016 (2001).
14. Watts, A. L. & Reddy, S. *arXiv:astro-ph/*0609364 (2006).
15. Page, D. & Cumming, A. *Astrophys. J. Lett.* **635,** L157–L160 (2005).

Competing financial interests: declared none.
doi:10.1038/nature05582

ASTROPHYSICS

Özel replies

Replying to: M. Alford *et al*. *Nature* 445, doi: 10.1038/nature05582 (2007)

In their comment[1] on my theoretical interpretation[2] of the observations of EXO 0748–676, Alford *et al*. suggest variants of quark-matter equations of state that produce stars consistent with my results. They do not challenge either the method that I propose or my conclusion that the data require a stiff equation of state[2].

Given the large uncertainties in the phenomenological description of quark matter used in predicting the properties of quark and hybrid stars, it is not surprising that models that meet the new constraints can be readily constructed. It is only through quantitative and uncoupled measurements of the masses and radii of neutron stars, such as the one that I propose[2], that the properties of matter at high densities will be constrained.

There are other methods related to the cooling of young neutron stars, pulsar glitches and quasiperiodic variability that offer the possibility of providing astrophysical constraints on the equation of state of the neutron-star matter. It is important to realize, however, that some of these methods rely on the two most uncertain properties of astrophysical objects, namely their distances and ages. Others lead to only qualitative inferences because of their strong dependence on models. In both cases, comparison of predictions with observations indicates that quark stars cannot be ruled out, but neither are they favoured[3–6].

Unlike these other methods, the one that I present results in a direct measurement of stellar masses and radii, with quantifiable uncertainties. In the case of EXO 0748–676, it leads to the firm conclusion that soft equations of state are ruled out, as Alford *et al*. concur.

F. Özel
Department of Physics, University of Arizona, Tucson, Arizona 85704, USA
e-mail: fozel@physics.arizona.edu

1. Alford, M. *et al. Nature* **445,** doi: 10.1038/nature05582 (2007).
2. Özel, F. *Nature* **441,** 1115–1117 (2006).
3. Kaplan, D. L. *et al. Astrophys. J.* **153** (suppl.)**,** 269–316 (2004).
4. Kaplan, D. L., Gaensler, B. M., Kulkarni, S. R. & Slane, P. O. *Astrophys. J.* **163** (suppl.)**,** 344–371 (2006).
5. Slane, P. O., Helfand, D. J. & Murray, S. S. *Astrophys. J. Lett.* **571,** L45–L49 (2002).
6. Watts, A. L. & Reddy, S. *arXiv:astro-ph/*0609364 (2006).

doi:10.1038/nature05583

© 2007 Nature Publishing Group

ARTICLES

Specialized hepatocyte-like cells regulate *Drosophila* lipid metabolism

Eugenio Gutierrez[1], David Wiggins[2], Barbara Fielding[2] & Alex P. Gould[1]

Lipid metabolism is essential for growth and generates much of the energy needed during periods of starvation. In *Drosophila*, fasting larvae release large quantities of lipid from the fat body but it is unclear how and where this is processed. Here we identify the oenocyte as the principal cell type accumulating lipid droplets during starvation. Tissue-specific manipulations of the Slimfast amino-acid channel, the Lsd2 fat-storage regulator and the Brummer lipase indicate that oenocytes act downstream of the fat body. In turn, oenocytes are required for depleting stored lipid from the fat body during fasting. Hence, lipid-metabolic coupling between the fat body and oenocytes is bidirectional. When food is plentiful, oenocytes have critical roles in regulating growth, development and feeding behaviour. In addition, they specifically express many different lipid-metabolizing proteins, including Cyp4g1, an ω-hydroxylase regulating triacylglycerol composition. These findings provide evidence that some lipid-processing functions of the mammalian liver are performed in insects by oenocytes.

In mammals, specialized cells of the adipose tissue and liver are critical for coordinating fat metabolism. This physiological axis regulates a complex set of lipid uptake, storage, synthesis, modification and degradation reactions essential for normal growth and development[1–3]. Lipid metabolism also has a particularly critical role in providing energy during periods of starvation. Food deprivation (fasting) stimulates the lipolysis of triglycerides (also called triacylglycerol, TAG) stored in adipocyte fat droplets via increases in hormone-sensitive lipase activity and adipose triglyceride lipase (ATGL) expression[4]. A large proportion of the fatty acids and other lipids thereby released into the circulation are then captured, modified and broken down by the major cell type of the liver, the hepatocyte[2,3]. An intriguing feature of the starvation response is that, in contrast to many other cell types, hepatocytes accumulate large numbers of fat droplets, resulting in hepatic steatosis. Fatty acids are released from hepatic lipid droplets during starvation and oxidized into shorter chain fatty acids and ultimately into soluble ketone bodies that can be discharged into the circulation for use as an energy source by many tissues. This fatty acid oxidation process involves chain shortening by α- and β-oxidation pathways active in peroxisomes and mitochondria[5,6]. Although lipid catabolism predominates during starvation, in the postprandial state, hepatocytes are highly active in synthesizing fatty acids for incorporation into triglycerides. These can then be assembled into lipoprotein particles, delivered to adipocytes and stored in lipid droplets[1]. One critical step for incorporating newly synthesized fatty acids into TAG is catalysed by stearoyl CoA-desaturase-1 (SCD-1), a hepatic enzyme converting palmitic ($C_{16:0}$) and stearic ($C_{18:0}$) acids into monounsaturated palmitoleic ($C_{16:1}$) and oleic ($C_{18:1}$) acids, respectively[7]. The importance of maintaining an appropriate balance between hepatic fatty acid synthesis and oxidation is highlighted by human diseases arising from mutations in fatty acid oxidation enzymes, and also by widespread diet-influenced pathologies such as non-alcoholic fatty liver disease and metabolic syndrome[8,9].

Invertebrate model organisms offer a powerful means to identify and functionally analyse lipid-metabolising genes. In *Caenorhabditis elegans*, fat is stored by intestinal epithelial cells and many regulators of this process have been identified using reverse genetic screens[10,11]. In contrast, *Drosophila* and other insects store lipid in a specialized tissue that resembles the adipose tissue of mammals, the fat body[12–14]. Diet-derived lipids, exported from the midgut as lipoproteins, are taken up from the haemolymph by the fat body via a mechanism involving Low-Density Lipoprotein (LDL) receptor-like molecules called Lipophorin receptors[15–17]. These lipids accumulate in fat body cells in the form of intracellular droplets but, when larvae are food-deprived, there is a net efflux of lipid into the haemolymph[18,19]. The mobilization process is regulated by TSC/TOR signalling and a nutrient sensor in the fat body that monitors amino-acid levels via the Slimfast (Slif) amino-acid channel[19]. Fasting-induced fat release is accompanied by increased lipolysis, at least in part associated with upregulation of Brummer, an ATGL-related lipase localized to lipid droplets[20]. Fat mobilization is also influenced by Lsd2, a lipid droplet protein related to a mammalian negative regulator of TAG hydrolysis called perilipin[21,22]. In addition to its involvement in lipid storage and release, the fat body produces a humoral signal regulating larval tissue growth in response to food availability[19,23–25]. Thus far, efforts to harness the power of *Drosophila* genetics to model human fat metabolism have been limited by the lack of information on how and where insect lipids are processed once they have been released from the fat body. For example, it is not known whether there is a specialized *Drosophila* tissue that synthesizes, modifies and oxidizes fatty acids in a similar way to the mammalian liver, nor is it clear to what degree the mammalian biochemical pathways metabolizing fatty acids are conserved in *Drosophila*. We now address both of these issues using a combination of bioinformatics, genetics and integrative physiology.

Oenocytes accumulate lipids during fasting

We first examined how fat is redistributed throughout the larval body after food deprivation. Using Oil Red O staining[26], three cell types in the third instar (L3) larva were found to contain numerous large (0.5–2.5 μm) lipid droplets under fed and/or fasting conditions: fat body cells, midgut epithelial cells and larval oenocytes (Fig. 1a). The fat body of L3 larvae has such a large capacity for lipid storage that,

[1]Medical Research Council, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK. [2]Oxford Centre for Diabetes, Endocrinology and Metabolism, Churchill Hospital, Oxford OX3 7LJ, UK.

©2007 Nature Publishing Group

despite lipid loss over a 14-h period of fasting, intense Oil Red O staining persists (100%, $n = 14$ fed larvae; 100%, $n = 24$ fasted larvae; Fig. 1b, c). As previously reported, we observe that lipid release from the L3 fat body during starvation correlates with lipid droplet aggregation[18,19] (also see insets to Fig. 1b, c). However, we find that droplet aggregation is not a reliable indicator of fasting at some other larval ages and durations of fasting (data not shown). In contrast to the fat body, regions of the gut (including the proventriculus and anterior midgut) staining strongly with Oil Red O under fed conditions (100%, $n = 14$) have only limited storage capacity, losing most lipid droplets after 14 h of fasting (100%, $n = 12$; Fig. 1d, e). The third cell type, larval oenocytes (distinct from adult oenocytes but abbreviated hereafter as oenocytes), are large cells of unknown function that are attached to the basal surface of the lateral epidermis in clusters of ~6 cells per abdominal hemi-segment[27–30]. We found that L3 oenocytes do not stain strongly with Oil Red O under fed conditions (100%, $n = 14$ larvae) but they do contain numerous large lipid droplets after a 14-h fast (96%, $n = 24$ larvae; Fig. 1f, g). This change in droplet abundance is consistent from oenocyte-to-oenocyte within one cluster and also from one cluster to another (Fig. 1h, i). Thus, oenocytes are highly atypical cells, in that they accumulate large numbers of lipid droplets specifically during fasting. As this is a hallmark of hepatocytes, the possibility is raised that insect oenocytes might process lipids in a similar way to the mammalian liver.

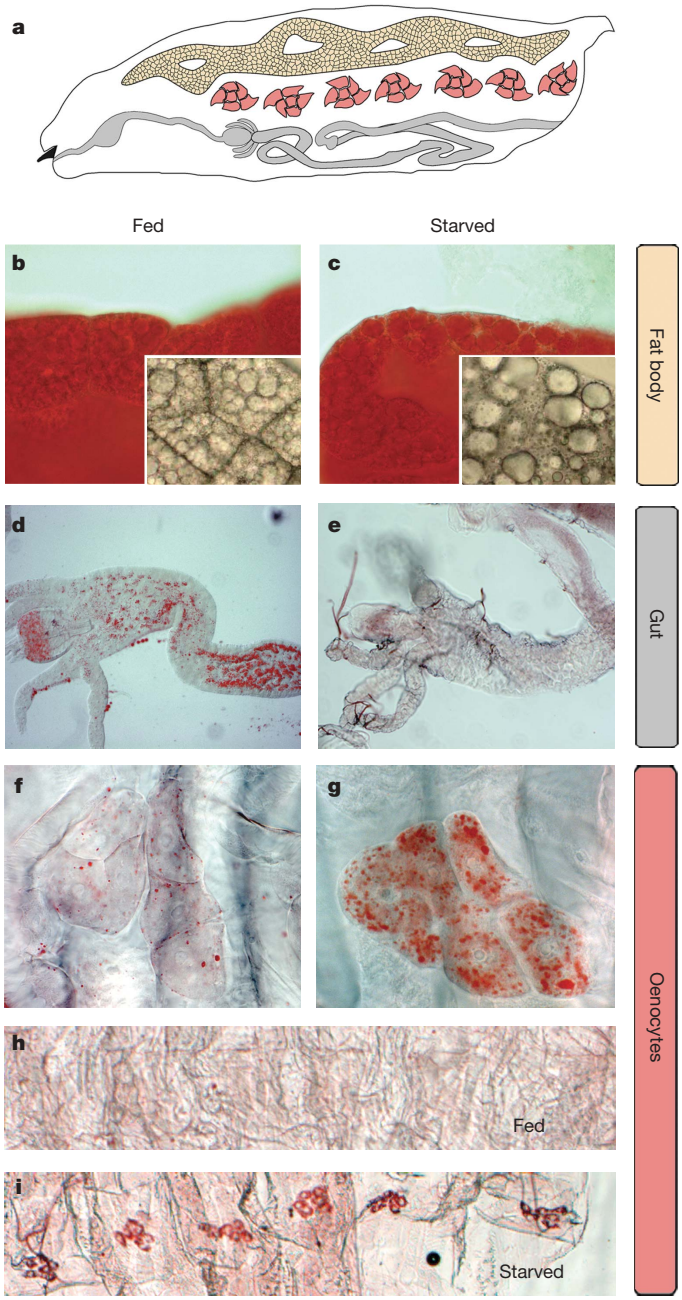


Figure 1 | Oenocytes accumulate lipids during starvation. **a**, Schematic drawing of a larva showing the three main tissues that accumulate lipid droplets: fat body (yellow), oenocytes (pink) and gut (grey). **b–i**, Oil Red O staining of L3 larvae under fed and starved conditions. After 14 h fasting, lipid in the fat body is not significantly depleted (**b**, **c**, insets show magnified images of unstained lipid droplets) but droplets in the proventriculus and anterior midgut are almost absent (**d**, **e**). **f–i**, Individual clusters of oenocytes (**f**, **g**) and larval pelts (**h**, **i**) showing that lipid droplets accumulate specifically and consistently in oenocytes after fasting.

The fat body regulates oenocyte lipid content

We next tested whether the accumulation of lipid droplets in oenocytes is regulated by the fat body nutrient sensor. An antisense transgene directed against the amino-acid transporter Slimfast was expressed in the fat body (*ppl-GAL4* driving *UAS-slif^{Anti}*; hereafter called *ppl>slif^{Anti}*). As reported previously[19], we observed that *ppl>slif^{Anti}* larvae raised to L3 on a standard diet (see Supplementary Methods) resemble starved wild-type larvae in that lipid droplets aggregate in the fat body (data not shown). Notably, we also found that oenocytes contain numerous lipid droplets, regardless of whether *ppl>slif^{Anti}* larvae are fed on a standard diet (100%, $n = 11$ larvae) or food-deprived for 14 h (88%, $n = 16$ larvae; Fig. 2a–c). This indicates that amino-acid monitoring via Slif in the fat body is required to ensure that lipid accumulation in oenocytes is kept low under standard nutritional conditions. TSC/TOR signalling, another component of the fat body nutrient sensor[19], is also involved, as overexpressing TSC1 and TSC2 (*ppl>Tsc1+2*) leads to a marked

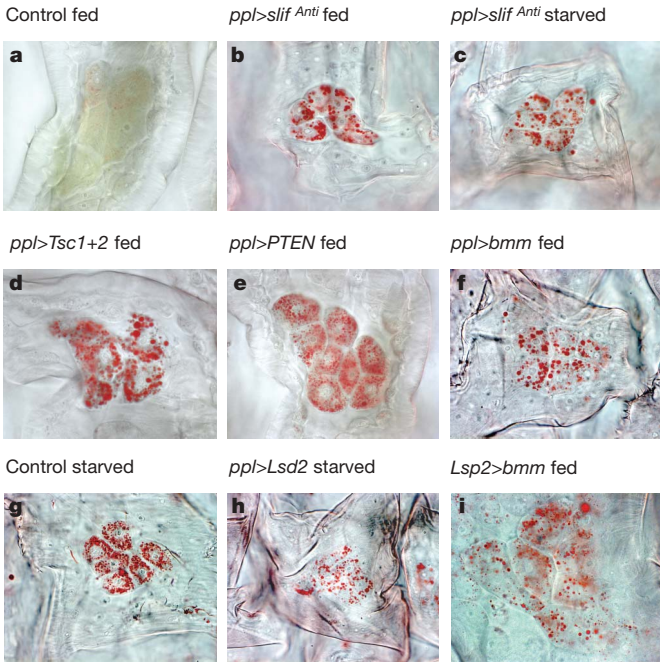


Figure 2 | The fat body regulates lipid accumulation in oenocytes. **a–c**, The oenocytes of *ppl>slif^{Anti}* L3 larvae contain many more Oil-Red-O-positive lipid droplets than fed controls (**a**), regardless of whether they are fed (**b**) or starved for 14 h (**c**). **d–f**, Oenocyte lipid droplets are numerous in L3 fed *ppl>Tsc1+2* (**d**), *ppl>PTEN* (**e**) and *ppl>bmm* (**f**) larvae. **g**, **h**, More oenocyte lipid droplets accumulate in starved control (**g**) than in starved *ppl>Lsd2* (**h**) L3 larvae. **i**, Oenocyte lipid droplets accumulate in fed *Lsp2>bmm* larvae at late L3.

©2007 Nature Publishing Group

accumulation of large lipid droplets in the oenocytes of 100% ($n = 11$) of fed larvae (Fig. 2d). Similarly, inhibiting the phosphatidylinositol-3 kinase pathway, which intersects with TOR signalling, by overexpressing the lipid phosphatase PTEN, also produces a build up of lipid droplets in the oenocytes of 89% ($n = 9$) of fed *ppl>PTEN* larvae (Fig. 2e). Hence, the fat body nutrient sensor regulates lipid accumulation in oenocytes but this could be directly via lipid release or indirectly, in response to a TSC/TOR-dependent signal.

To assess directly the effect of lipid mobilization from the fat body, the balance between TAG storage and hydrolysis was altered in two ways. First, we overexpressed Brummer (Bmm) lipase, which is normally limiting for lipid release from the fat body[20]. This is sufficient to produce specific accumulation of lipid droplets in the oenocytes of 92% ($n = 12$) of fed *ppl>bmm* larvae (Fig. 2f). Second, TAG release from lipid droplets was decreased by overexpressing Lsd2. This reduces the accumulation of oenocyte lipid droplets in 78% ($n = 9$) of starved *ppl>Lsd2* larvae, with ~4-fold fewer large droplets per oenocyte (Fig. 2g, h and data not shown). We also used a second driver, *Lsp2-GAL4*, that unlike *ppl-GAL4* is activated in the fat body only at the mid-L3 stage[31]. This temporally restricted driver nevertheless suffices to induce oenocyte lipid droplet accumulation in 100% ($n = 5$) of fed *Lsp2>bmm* larvae (Fig. 2i) and also in fed *Lsp2>slif^{Anti}* animals (data not shown). Together, the Slif, TSC, PTEN, Lsd2 and Bmm results demonstrate metabolic regulation from the fat body to the oenocytes, although they do not exclude the involvement of intermediate tissues such as the gut. Either way, these results strongly suggest that, when nutrition is poor, falling amino-acid levels stimulate lipid release from the fat body and subsequent lipid uptake from the haemolymph by oenocytes.

Oenocytes regulate larval growth and feeding

To determine the *in vivo* roles of oenocytes during fasting and normal development, we developed a targeted binary cell ablation system. Larvae carrying *sal[BO,7.6kb]GAL4*, a purpose-built oenocyte driver, and also *UAS-reaper*, an inducible pro-apoptotic transgene, lack 100% of oenocytes from L1 onwards and die before reaching pupariation (hereafter called *BO>rpr* larvae). As a specificity control, *BO>rpr* animals were rescued to viable adults by expressing Gal80, an inhibitor of Gal4, under the regulation of an independent oenocyte enhancer from *seven up* (*svp*). As *svp[3kb]GAL80* suppresses *sal[BO,7.6kb]GAL4* activity in oenocytes but not in secondary larval sites, *BO>rpr* lethality results from the ablation of oenocytes and not some other cell type (Supplementary Fig. 1).

BO>rpr larvae raised on a standard diet attain a similar mass to *UAS-rpr* controls during L1 but, after the L1-to-L2 transition, they grow at a much slower rate (Fig. 3a). Notably, reduced growth correlates with aberrant feeding behaviour, with most *BO>rpr* larvae dispersing away from the yeast food source during L2 (Supplementary Fig. 2a). This dispersal is distinct from premature wandering behaviour as *BO>rpr* larvae enter and exit the yeast source multiple times, retain food in the gut and do not pupariate precociously (Fig. 3c and data not shown). As *BO>rpr* larvae spend less time in the food source and grow more slowly than L2 controls, we examined whether they show increased mouth-hook contractions, a behavioural response to hunger[32]. However, we observed reduced mouth-hook contractions that are not significantly increased by the motivation of a 2-h period of food deprivation (Supplementary Fig. 2b). Thus, rather than stimulating hunger-driven feeding behaviour, oenocyte ablation seems to block it, although this effect could be very indirect. Either way, reduced feeding is likely to contribute to the slow growth rate of *BO>rpr* larvae during L2.

Oenocytes regulate developmental progression

As reduced growth resulting from inadequate nutrition before 70 h after egg laying (just before the L2/L3 moult) is associated with larval arrest rather than smaller-than-normal adult flies[33], we used morphological criteria to stage oenocyte-ablated animals. We observed that *BO>rpr* larvae arrest at several different stages after the L1/L2 transition, thus displaying a polyphasic lethality profile (Fig. 3b and Supplementary Fig. 3a, b). Although arrested development can result from reduced signalling by ecdysteroids[34,35], the *BO>rpr* polyphasic lethality profile is not significantly altered by adding 20-hydroxyecdysone or its precursor ecdysone (Supplementary Fig. 3c, d). Therefore, a deficiency in these ecdysteroids is not the sole reason for *BO>rpr* arrest, but we can not exclude the possibility that it, together with some other oenocyte deficit, contributes to the moulting phenotype.

Unlike many larval tissues, oenocytes persist for much of pupal development (Fig. 3d–f). To address whether oenocytes are required for metamorphosis, we used a temperature-sensitive version of Gal80 (*GAL80^{ts}*, ref. 36) to attenuate Gal4 activity, thus bypassing *BO>rpr* larval lethality. Combining *tub-GAL80^{ts}* with *BO>rpr* suppresses apoptosis in approximately 50% of oenocytes at 25 °C (from L1 onwards) and allows developmental progression until pupal stages. However, no animals complete pupal development, with many failing to separate from the puparial case during eclosion (Fig. 3g, h).

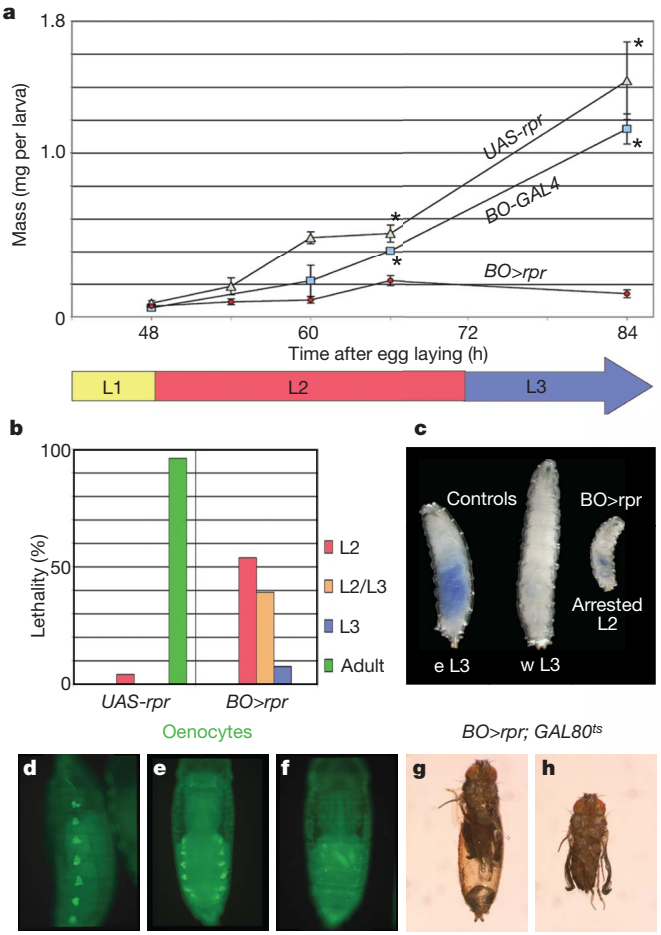


Figure 3 | Oenocytes regulate growth and developmental progression. **a**, *BO>rpr* larvae, lacking oenocytes, grow less than *UAS-rpr* and *BO-GAL4* controls during L2 and L3. Error bars (±1 s.d.), significance at the $P < 0.01$ level (asterisk) and hours after egg laying are indicated. **b**, **c**, *BO>rpr* animals show polyphasic larval lethality (**b**) and developmental arrest in L2 without gut clearance of blue-labelled yeast (**c**). In panel **c**, control L3 larvae are shown at early (eL3) and late/wandering (wL3) stages. **d**–**f**, Oenocytes (GFP labelled) retain normal morphology in the pre-pupa (**d**), become incorporated into the pharate adult abdomen (**e**) and then fragment (**f**). **g**, **h**, Ablation of ~50% oenocytes in *BO>rpr; GAL80^{ts}* animals leads to pupal lethality with frequent arrest during adult eclosion, visualized before (**g**) and after (**h**) dissection from the puparial case.

©2007 Nature Publishing Group

Together, the 50% and 100% oenocyte ablation phenotypes demonstrate that oenocytes are required for growth and developmental progression during both larval and pupal stages.

Oenocytes regulate lipid content

We next examined whether lipid metabolism is altered in larvae lacking all oenocytes. At early L2, when *BO>rpr* larvae are the same size as controls, no significant abnormalities in TAG content or in the relative amounts of the major long-chain fatty acids were detected (Fig. 4a and data not shown). The fat storage capacity of larvae at early L2 is much less than at L3 such that a 12-h period of food withdrawal is sufficient to deplete ~60% of stored TAG in control animals (Fig. 4a). However, during this same fasting period, *BO>rpr* larvae only lose ~10% of total TAG (Fig. 4a). This deficit in TAG depletion correlates with a higher density of fat-body lipid droplets in 100% of starved *BO>rpr* larvae ($n = 9$) compared to controls ($n = 7$) after fasting (Fig. 4b–e). As ~80% of larval fatty acids are stored as TAG, we also examined whether the proportions of individual fatty acids are altered in fasting *BO>rpr* larvae. In early L2 controls, lauric ($C_{12:0}$) and myristic ($C_{14:0}$) acids are depleted more efficiently than longer-chain (C_{16}–C_{20}) fatty acids such that their mass, relative to stearic acid ($C_{18:0}$), is reduced twofold after 12 h fasting. However, in fasted *BO>rpr* larvae, the $C_{12:0}/C_{18:0}$ and $C_{14:0}/C_{18:0}$ ratios remain close to those before starvation, corresponding to approximately twice the value of starved controls (Fig. 4f). Together, these results indicate that oenocytes are required for efficiently depleting fatty acids, stored largely in the fat body as TAG, during nutrient deprivation. With the previous Slif, TSC, PTEN, Bmm and Lsd2 results, we propose that lipid-metabolic coupling between the fat body and oenocytes is bidirectional.

Oenocytes express many lipid metabolizing genes

To identify the metabolic pathways processing lipids within oenocytes, we first identified 51 genes expressed selectively or exclusively in oenocytes (see Supplementary Methods). About 40% of these (22 out of 51) encode orthologues of known human lipid-metabolizing/processing proteins (Supplementary Table 1, Supplementary Fig. 4 and data not shown). The high degree of conservation of most *Drosophila* proteins, together with some previous functional studies, suggests that oenocytes express lipid metabolic pathways strikingly similar to those of hepatocytes. By analogy, oenocytes would capture lipid from lipophorin in the haemolymph via LpR1 and LpR2, two Lipophorin receptors. Fatty acids released from lipid droplets by lipases such as the *CG17292* product, could then be modified by a variety of enzymes, including the Desat1 and CG9743 acyl-CoA desaturases[37], the CG18609 and CG6921 fatty acid elongases and the microsomal lipid ω-hydroxylase, Cytochrome P450-4g1 (Cyp4g1)[38]. Fatty acids could also be chain shortened, at least partially, by the actions of peroxisomal β-oxidation components including those encoded by *CG11151* (similar to *Sterol carrier protein 2*), *CG12428* (*Carnitine O-octanoyl transferase*), *CG9527* (*Pristanoyl-CoA oxidase*) and *Catalase*, the peroxisomal enzyme inactivating oxygen free radicals produced by pristanoyl-CoA oxidases[39]. In addition, oenocytes strongly express Hnf4 and Svp, orthologues of the mammalian nuclear receptors Hnf4-α and COUP-TF, known regulators of hepatocyte differentiation and lipid-metabolic genes. Thus, the oenocyte/hepatocyte analogy includes a shared set of lipid-metabolizing genes and at least two of their transcriptional regulators.

Cyp4g1 regulates TAG composition

To explore the functions of fatty acid metabolism specifically within *Drosophila* oenocytes, we generated two lethal protein-null alleles for the predicted lipid ω-hydroxylase encoded by *Cyp4g1* (Fig. 5a). *Cyp4g1* is known to be expressed in oenocytes[30,40,41] and we now find that this is the only site of detectable expression in embryos and larvae (Fig. 5b and Supplementary Fig. 4). Animals homozygous for either the $Cyp4g1^{\Delta 4}$ or $Cyp4g1^{\Delta 14\text{-}9}$ allele develop normally through larval and early pupal stages but arrest during mid-to-late pupal stages, with many failing during adult eclosion (Fig. 5c). This pupal phenotype is strikingly similar to the 50% oenocyte ablation phenotype. Moreover, although late-L3 *Cyp4g1* mutant larvae appear morphologically indistinguishable from controls, they manifest a twofold increase in the oleic acid:stearic acid ratio ($C_{18:1}/C_{18:0}$). Notably, this imbalance in fatty acid desaturation is found in the TAG fraction but not in the phospholipid fraction (Fig. 5d). This selectivity strongly suggests that the *Cyp4g1* defect is specific to fatty acids in metabolic storage form, most of which reside in the fat body, rather than fatty acids present in the structural lipids of all cell membranes. Taken together, the metabolic profiles of oenocyte-ablated and *Cyp4g1* mutant larvae provide two independent lines of evidence that oenocytes regulate the lipid composition of the fat body.

Discussion

Larval oenocytes were first described in insects over 140 yr ago[42]. Since then, however, their functions have remained unclear, with largely descriptive studies implicating them in processes such as cuticle synthesis and the regulation of haemolymph composition[43–45]. Using cell ablation to test their functions directly for the first time, we find clear requirements for larval growth and pupal development. Although the subset of oenocyte genes mediating the larval developmental functions remains to be identified, for pupal development we have shown that the lipid ω-hydroxylase *Cyp4g1* is required. At least one important role of *Cyp4g1* is to downregulate the ratio of oleic-to-stearic acid, widely used as a marker of SCD-1 activity in mammals. This prompts speculation that Cyp4g1 may repress the activity of stearoyl CoA-desaturases like Desat1 (see Supplementary

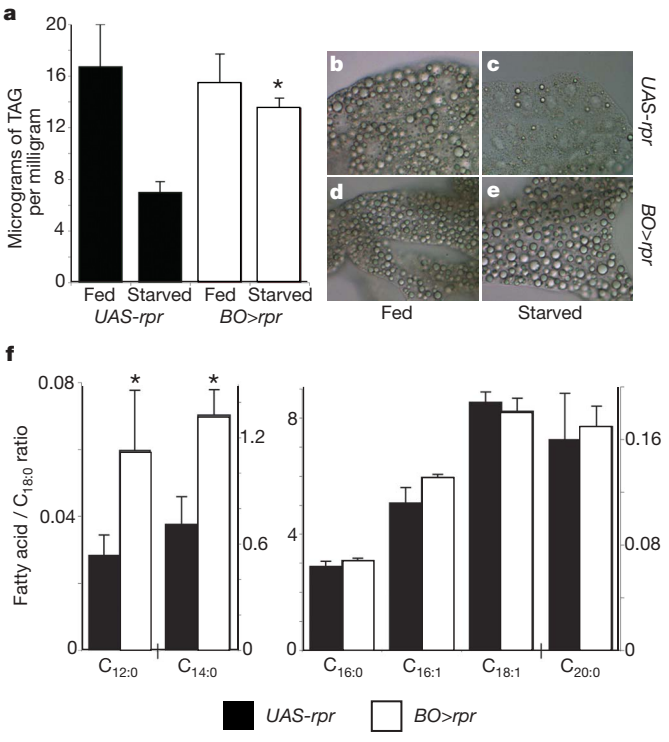


Figure 4 | Oenocytes are required for lipid depletion from the fat body during starvation. a, TAG content of *UAS-rpr* (control) and *BO>rpr* larvae before (fed) and after (starved) fasting at early L2. Asterisk, $P < 0.001$ for starved *BO>rpr* versus starved *UAS-rpr* controls. In this, and all subsequent histograms, error bars represent 1 s.d. from the mean of triplicate measurements. **b**–**e**, Lipid droplets in unstained fat body of *UAS-rpr* and *BO>rpr* larvae before and after starvation. **f**, Composite histograms showing ratios of individual fatty acids, relative to stearic acid ($C_{18:0}$), in starved *UAS-rpr* and *BO>rpr* larvae (see text). The proportion of $C_{18:0}$ does not change significantly. Asterisk, $P < 0.01$ for $C_{14:0}$ and $P < 0.05$ for $C_{12:0}$. Note that there are four different *y*-axis scales for $C_{12:0}$, $C_{14:0}$, $C_{16:0}$-$C_{18:1}$ and $C_{20:0}$.

©2007 Nature Publishing Group

Fig. 4), thereby inhibiting inappropriate monounsaturated fatty acid synthesis during long non-feeding periods such as late L3 and pupal stages.

Four lines of evidence argue that at least some of the lipid-metabolizing roles of insect oenocytes are analogous to those of mammalian hepatocytes. First, oenocytes express 22 orthologues of human fat-metabolizing genes expressed in hepatocytes. Second, like hepatocytes, they are atypical cells in that they accumulate fat droplets during starvation. Third, like the liver, oenocytes lie downstream of a nutrient sensor in a major fat depot. And fourth, Brummer lipase and Lsd2 in the fat body regulate oenocyte lipid content in a broadly similar way as ATGL and perilipin in adipose tissue regulate hepatic fat influx. However, whereas hepatocytes store large quantities of glycogen, this role in *Drosophila* is primarily carried out by the fat body[14]. Thus, mammalian liver functions in glycogen storage and lipid processing seem to be divided in *Drosophila* between the fat body and oenocytes.

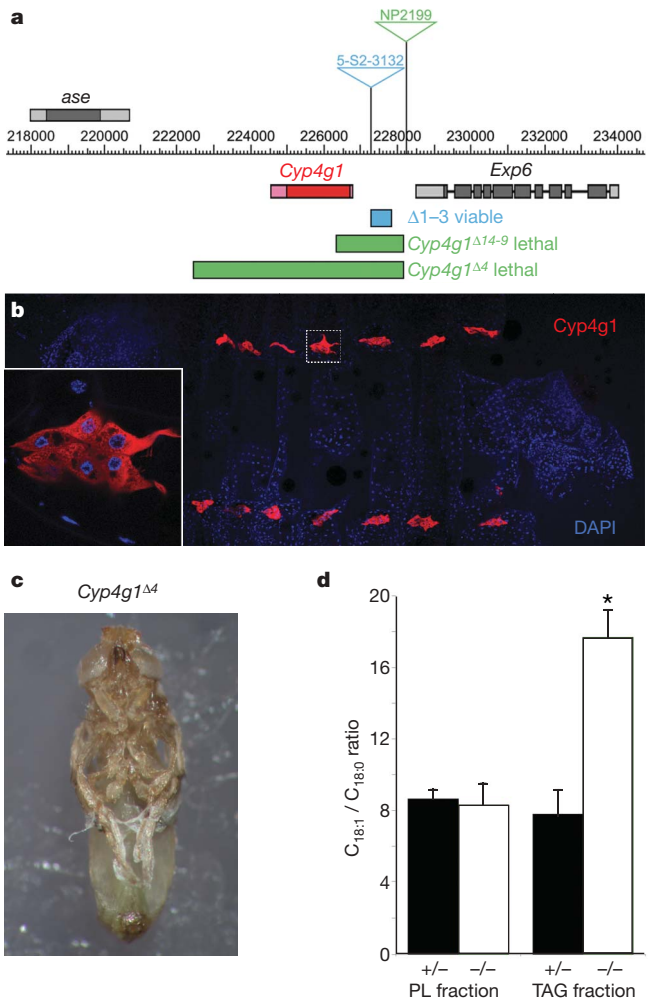


Figure 5 | *Cyp4g1* is an essential oenocyte-specific gene regulating TAG composition. a, *Cyp4g1* locus depicting P-element insertions and viable (blue) and lethal (green) deletions. Predicted translated (dark fill) and untranslated (light fill) regions of *Cyp4g1*, *asense* (*ase*) and *Exportin 6* (*Exp6*), and a scale bar (release 3.1 coordinates) are shown. **b**, Larval pelt showing that Cyp4g1 expression is confined to oenocyte clusters. Inset of boxed region shows one cluster. **c**, $Cyp4g1^{\Delta 4}$ hemizygote arrested during adult eclosion, shown dissected from the puparial case. **d**, Histogram showing elevated oleic acid:stearic acid ($C_{18:1}/C_{18:0}$) ratio in L3 larvae that are $Cyp4g1^{\Delta 14\text{-}9}$ mutant (−/−) compared to *FM7c, Kr>GFP* balanced controls (+/−). Asterisk, $P < 0.002$ for −/− versus +/− TAG fractions. PL, phospholipid.

This study suggests the existence of two-way metabolic coupling between the fat body and oenocytes. Analogous to the mammalian adipose–liver axis, lipid mobilization from the fat body during starvation produces lipid droplet accumulation in oenocytes, a metabolic change resembling hepatic steatosis. In *Drosophila*, we also found a reciprocal regulation, namely that oenocytes are required for efficiently depleting lipid from the fat body during fasting. This suggests a feedback mechanism for matching lipid supply to demand, whereby oenocytes keep haemolymph lipids low and also promote lipid mobilization from the fat body. Thus, in oenocyte-less larvae, excess circulating lipids might underlie the behavioural syndrome of larval dispersal and reduced feeding in a similar way as reported for elevated amino-acid levels[46]. Central to our proposed feedback model is the signal acting on the lipogenesis/lipolysis balance within the fat body. The data presented here are equally compatible with this signal corresponding to a haemolymph lipid/metabolite or to a separate signal generated by oenocytes. Regarding the latter possibility, it is interesting that recent work in mammals indicates that the liver secretes signalling factors (hepatokines) that promote lipolysis in adipose tissue[47]. Together with our study, this suggests that *Drosophila* may prove useful, not only for modelling hepatic steatosis, but also some regulatory roles of the liver in metabolic homeostasis.

METHODS

***Drosophila* analysis.** See Supplementary Methods for full details of bioinformatic analysis, *Drosophila* strains, generation of *Cyp4g1* mutants, construction of *BO-GAL4* and *svp-GAL80* transgenes, larval fasting, and other analyses.

Oil Red O staining. Larvae were partially dissected in PBS to expose the inner face of the epidermis and fixed in 4% paraformaldehyde/PBS for 10 min. Specimens were then transferred to netted baskets, rinsed twice with distilled water, incubated for 20 to 30 min in Oil Red O stain (6 ml of 0.1% Oil Red O in isopropanol and 4 ml distilled water: prepared fresh and passed through a 0.45-μm syringe filter), blotted onto tissue paper and rinsed twice with distilled water. Stained specimens were mounted in glycerol and scored for large Oil-Red-O-positive droplets.

Lipid analysis. Total lipid extracts and individual lipid fractions separated by thin-layer chromatography were prepared from larvae ground to a fine powder in liquid nitrogen, as described for rodent liver[48]. TAG was determined enzymatically from total lipid extracts using a GPO-PAP triglyceride kit (Randox Laboratories). Fatty acid profiles were determined from trans-methylated total or fractionated lipids[49] by gas chromatography using a 30 m × 0.53 mm × 1 μm HP-FFAP-123 capillary column (Agilent Technologies). Individual fatty acids were quantified by inclusion of known quantities of tri-pentadecanoin. *P*-values were calculated, assuming equal sample variance, using two-tailed Student's *t*-tests.

Received 22 July; accepted 23 October 2006.
Published online 29 November 2006.

1. Gibbons, G. F., Islam, K. & Pease, R. J. Mobilisation of triacylglycerol stores. *Biochim. Biophys. Acta* **1483**, 37–57 (2000).
2. Lee, C. H., Olson, P. & Evans, R. M. Minireview: lipid metabolism, metabolic diseases, and peroxisome proliferator-activated receptors. *Endocrinology* **144**, 2201–2207 (2003).
3. Yu, S., Rao, S. & Reddy, J. K. Peroxisome proliferator-activated receptors, fatty acid oxidation, steatohepatitis and hepatocarcinogenesis. *Curr. Mol. Med.* **3**, 561–572 (2003).
4. Zechner, R., Strauss, J. G., Haemmerle, G., Lass, A. & Zimmermann, R. Lipolysis: pathway under construction. *Curr. Opin. Lipidol.* **16**, 333–340 (2005).
5. Reddy, J. K. & Hashimoto, T. Peroxisomal β-oxidation and peroxisome proliferator-activated receptor-α: an adaptive metabolic system. *Annu. Rev. Nutr.* **21**, 193–230 (2001).
6. Wanders, R. J. Peroxisomes, lipid metabolism, and peroxisomal disorders. *Mol. Genet. Metab.* **83**, 16–27 (2004).
7. Ntambi, J. M. & Miyazaki, M. Regulation of stearoyl-CoA desaturases and role in metabolism. *Prog. Lipid Res.* **43**, 91–104 (2004).
8. Wanders, R. J. *et al.* Peroxisomal fatty acid α- and β-oxidation in health and disease: new insights. *Adv. Exp. Med. Biol.* **544**, 293–302 (2003).
9. Browning, J. D. & Horton, J. D. Molecular mediators of hepatic steatosis and liver injury. *J. Clin. Invest.* **114**, 147–152 (2004).
10. McKay, R. M., McKay, J. P., Avery, L. & Graff, J. M. C. *elegans*: a model for exploring the genetics of fat storage. *Dev. Cell* **4**, 131–142 (2003).
11. Ashrafi, K. *et al.* Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* **421**, 268–272 (2003).

©2007 Nature Publishing Group

12. Butterworth, F. M., Bodenstein, D. & King, R. C. Adipose tissue of *Drosophila melanogaster*. I. An experimental study of larval fat body. *J. Exp. Zool.* **158,** 141–153 (1965).
13. Keeley, L. L. in *Comprehensive Insect Physiology, Biochemistry and Pharmacology* (eds Kerkut, G. A. & Gilbert, L. I.) 211–248 (Pergamonn, New York, 1985).
14. Dean, R. L., Locke, M. & Collins, J. V. in *Comprehensive Insect Physiology, Biochemistry and Pharmacology* (eds Kerkut, G. A. & Gilbert, L. I.) 155–210 (Pergamonn, New York, 1985).
15. Canavoso, L. E., Jouni, Z. E., Karnas, K. J., Pennington, J. E. & Wells, M. A. Fat metabolism in insects. *Annu. Rev. Nutr.* **21,** 23–46 (2001).
16. Dantuma, N. P. *et al.* An insect homolog of the vertebrate very low density lipoprotein receptor mediates endocytosis of lipophorins. *J. Lipid Res.* **40,** 973–978 (1999).
17. Lee, C. S. *et al.* Wax moth, *Galleria mellonella*, high density lipophorin receptor: alternative splicing, tissue-specific expression, and developmental regulation. *Insect Biochem. Mol. Biol.* **33,** 761–771 (2003).
18. Zhang, H., Stallock, J. P., Ng, J. C., Reinhard, C. & Neufeld, T. P. Regulation of cellular growth by the *Drosophila* target of rapamycin dTOR. *Genes Dev.* **14,** 2712–2724 (2000).
19. Colombani, J. *et al.* A nutrient sensor mechanism controls *Drosophila* growth. *Cell* **114,** 739–749 (2003).
20. Gronke, S. *et al.* Brummer lipase is an evolutionary conserved fat storage regulator in *Drosophila. Cell Metab.* **1,** 323–330 (2005).
21. Gronke, S. *et al.* Control of fat storage by a *Drosophila* PAT domain protein. *Curr. Biol.* **13,** 603–606 (2003).
22. Teixeira, L., Rabouille, C., Rorth, P., Ephrussi, A. & Vanzo, N. F. *Drosophila* Perilipin/ADRP homologue Lsd2 regulates lipid metabolism. *Mech. Dev.* **120,** 1071–1081 (2003).
23. Britton, J. S. & Edgar, B. A. Environmental control of the cell cycle in *Drosophila*: nutrition activates mitotic and endoreplicative cells by distinct mechanisms. *Development* **125,** 2149–2158 (1998).
24. Martin, J. F., Hersperger, E., Simcox, A. & Shearn, A. minidiscs encodes a putative amino acid transporter subunit required non-autonomously for imaginal cell proliferation. *Mech. Dev.* **92,** 155–167 (2000).
25. Britton, J. S., Lockwood, W. K., Li, L., Cohen, S. M. & Edgar, B. A. *Drosophila*'s insulin/PI3-kinase pathway coordinates cellular metabolism with nutritional conditions. *Dev. Cell* **2,** 239–249 (2002).
26. Lillie, R. D. Various oil soluble dyes as fat stains in the supersaturated isopropanol technic. *Stain Technol.* **19,** 55–58 (1944).
27. Lawrence, P. A. & Johnston, P. Observations on cell lineage of internal organs of *Drosophila. J. Embryol. Exp. Morphol.* **91,** 251–266 (1986).
28. Elstob, P. R., Brodu, V. & Gould, A. P. *spalt*-dependent switching between two cell fates that are induced by the *Drosophila* EGF receptor. *Development* **128,** 723–732 (2001).
29. Rusten, T. E. *et al.* Spalt restricts EGFR mediated induction of chordotonal precursors in the embryonic PNS of *Drosophila. Development* **128,** 711–722 (2001).
30. Brodu, V., Elstob, P. R. & Gould, A. P. EGF receptor signaling regulates pulses of cell delamination from the *Drosophila* ectoderm. *Dev. Cell* **7,** 885–895 (2004).
31. Cherbas, L., Hu, X., Zhimulev, I., Belyaeva, E. & Cherbas, P. EcR isoforms in *Drosophila*: testing tissue-specific requirements by targeted blockade and rescue. *Development* **130,** 271–284 (2003).
32. Wu, Q., Zhang, Y., Xu, J. & Shen, P. Regulation of hunger-driven behaviors by neural ribosomal S6 kinase in *Drosophila. Proc. Natl Acad. Sci. USA* **102,** 13289–13294 (2005).
33. Beadle, G. W., Tatum, E. L. & Clancy, C. W. Food level in relation to rate of development and eye pigmentation in *Drosophila. Biol. Bull.* **75,** 447–462 (1938).
34. Riddiford, L. M. in The Development of *Drosophila melanogaster* (eds Bate, M. & Martinez Arias, A.) 899–939 (Cold Spring Harbor Laboratory Press, New York, 1993).
35. Gilbert, L. I. Halloween genes encode P450 enzymes that mediate steroid hormone biosynthesis in *Drosophila melanogaster. Mol. Cell. Endocrinol.* **215,** 1–10 (2004).
36. McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K. & Davis, R. L. Spatiotemporal rescue of memory dysfunction in *Drosophila. Science* **302,** 1765–1768 (2003).
37. Ueyama, M., Chertemps, T., Labeur, C. & Wicker-Thomas, C. Mutations in the *desat1* gene reduces the production of courtship stimulatory pheromones through a marked effect on fatty acids in *Drosophila melanogaster. Insect Biochem. Mol. Biol.* **35,** 911–920 (2005).
38. Sundseth, S. S., Nix, C. E. & Waters, L. C. Isolation of insecticide resistance-related forms of cytochrome P-450 from *Drosophila melanogaster. Biochem. J.* **265,** 213–217 (1990).
39. Griswold, C. M., Matthews, A. L., Bewley, K. E. & Mahaffey, J. W. Molecular characterization and rescue of acatalasemic mutants of *Drosophila melanogaster. Genetics* **134,** 781–788 (1993).
40. Simpson, A. E. The cytochrome P450 4 (CYP4) family. *Gen. Pharmacol.* **28,** 351–359 (1997).
41. Tomancak, P. *et al.* Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome Biol.* **3,** RESEARCH0088 (2002).
42. Landois, L. Ueber die Funktion des Fettkörpers. *Zeitschr. f. Wissensch. Zoologie* **15,** 371–372 (1865).
43. Koschevnikov, G. Ueber den Fettkörper und die Oenocyten der Honigbiene (*Apis mellifera*). *Z. Anz. Bd.* **23,** 657–661 (1900).
44. Koller, G. Die innere Sekretion bei wirbellosen Tieren. *Biol. Rev.* **4,** 269–306 (1928).
45. Wigglesworth, V. The physiology of the cuticle and of ecdysis in *Rhodnius prolixus* (Triatomidae, Hemiptera); with special reference to the function of the oenocytes and of the dermal glands. *Quart. J. Micros. Sci.* **76,** 269–318 (1933).
46. Zinke, I., Kirchner, C., Chao, L. C., Tetzlaff, M. T. & Pankratz, M. J. Suppression of food intake and growth by amino acids in *Drosophila*: the role of *pumpless*, a fat body expressed gene with homology to vertebrate glycine cleavage system. *Development* **126,** 5275–5284 (1999).
47. Oike, Y., Akao, M., Kubota, Y. & Suda, T. Angiopoietin-like proteins: potential new targets for metabolic syndrome therapy. *Trends Mol. Med.* **11,** 473–479 (2005).
48. Duerden, J. M. & Gibbons, G. F. Secretion and storage of newly synthesized hepatic triacylglycerol fatty acids *in vivo* in different nutritional states and in diabetes. *Biochem. J.* **255,** 929–935 (1988).
49. Burdge, G. C., Wright, P., Jones, A. E. & Wootton, S. A. A method for separation of phosphatidylcholine, triacylglycerol, non-esterified fatty acids and cholesterol esters from plasma by solid-phase extraction. *Br. J. Nutr.* **84,** 781–787 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Gibbons and K. Frayn for providing advice and facilities for metabolic profiling, and P. Elstob, V. Brodu, S. Mahadevaiah and P. Sarchet for assistance with enhancer trap screening and sequencing. We also thank S. Celniker and BDGP for providing many of the panels used in Supplementary Fig. 4, and R. Barrio, B. Bello, L. Michaut, A. Brand, R. Carthew, W. Janning, S. Kennel, C. O'Kane, R. Kühnlein, P. Leopold, I. Miguel-Aliaga, I. Salecker, R. Schultz, N. Tapon, C. Thummel, J.-P. Vincent, T. Xu, Flyview, The NP consortium, the DGRC at Kyoto Institute of Technology, the Bloomington, Umeå, and Szeged stock centres and the DSHB at the University of Iowa for DNA constructs, flies and antibodies. We also thank I. Robinson for discussions and J. Briscoe, X. Franch-Marro, G. Gibbons, I. Miguel-Aliaga, E. Ober, E. Piddini, I. Salecker, P. Serpente and D. Wilkinson for critical reading of the manuscript. This work was supported by the Medical Research Council (E.G., D.W and A.P.G.), the Mexican National Council for Science and Technology (E.G.) and the Wellcome Trust (B.F.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.P.G. (agould@nimr.mrc.ac.uk).

©2007 Nature Publishing Group

ARTICLES

Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast

Philip Zegerman[1] & John F. X. Diffley[1]

Cyclin-dependent kinases (CDKs) drive major cell cycle events including the initiation of chromosomal DNA replication. We identified two S phase CDK (S-CDK) phosphorylation sites in the budding yeast Sld3 protein that, together, are essential for DNA replication. Here we show that, when phosphorylated, these sites bind to the amino-terminal BRCT repeats of Dpb11. An Sld3–Dpb11 fusion construct bypasses the requirement for both Sld3 phosphorylation and the N-terminal BRCT repeats of Dpb11. Co-expression of this fusion with a phospho-mimicking mutant in a second essential CDK substrate, Sld2, promotes DNA replication in the absence of S-CDK. Therefore, Sld2 and Sld3 are the minimal set of S-CDK targets required for DNA replication. DNA replication in cells lacking G1 phase CDK (G1-CDK) required expression of the Cdc7 kinase regulatory subunit, Dbf4, as well as Sld2 and Sld3 bypass. Our results help to explain how G1- and S-CDKs promote DNA replication in yeast.

DNA replication in eukaryotic cells is regulated by a conserved two-step mechanism[1,2]. In the first step (licensing), the putative replication fork helicase Mcm2–7 is loaded into pre-replicative complexes. CDKs directly prevent pre-replicative complex assembly in the budding yeast *Saccharomyces cerevisiae* by a variety of well-characterized mechanisms[2]. Consequently, pre-replicative complexes can assemble only during G1 phase, when CDK levels are low. In the second step, CDKs together with a second essential protein kinase, Dbf4/Cdc7, promote initiation[3–8]. Thus, oscillation of CDK activity during the cell cycle ensures that replication origins fire just once during S phase.

Although proteomic approaches have identified nearly 200 CDK substrates in yeast[9], to date only one S-CDK substrate—Sld2/Drc1 (refs 10, 11)—has been identified that requires phosphorylation for initiating DNA replication[12,13]. Phospho-mimicking Sld2 mutants do not bypass the requirement for S-CDK[12] indicating that other essential S-CDK substrates exist, but the number and identity of these substrates are unknown. Here we show that another replication initiation factor, Sld3, is a second essential S-CDK substrate. Moreover, using bypass mutants we show that Sld2 and Sld3 represent the minimal set of S-CDK substrates required for DNA replication.

Sld3 is an essential CDK substrate

While examining potential S-phase-specific protein kinase substrates, we constructed an allele of *SLD3* encoding three Ser/Thr to Ala mutations (T600, T609 and S622), which was lethal (Fig. 1a). Although all three single mutants were viable, *sld3-600A* had a pronounced slow growth phenotype, particularly at 24 °C (Fig. 1b). The double-mutant *sld3-600A,622A* was lethal whereas *sld3-600A,609A* was viable but grew slowly, similar to the *sld3-600A* single mutant. When *sld3-600A* cells were released from a block in G1 phase with α factor, S phase proceeded very slowly relative to wild type (Fig. 1c), indicating an important role for T600 in DNA replication. Mutation of T600, T609 and S622 to aspartic acid did not overcome the essential requirement for these sites, suggesting that they do not act as phospho-mimics (Fig. 1b).

T600, T609 and S622 are all followed by proline residues, the minimal sequence required for CDK phosphorylation and can serve as substrates for purified S-CDK *in vitro* (Supplementary Fig. 1a). Additionally, *sld3-600A* is synthetically lethal with deletion of the S phase cyclin Clb5 (Supplementary Fig. 1b), suggesting interactions between Sld3 and S-CDK. We generated a polyclonal antibody that specifically recognized phospho-T600 (Supplementary Fig. 2), the most important site (Fig. 1b), and used this antibody to examine Sld3 phosphorylation *in vivo*. T600 was hypo-phosphorylated in α-factor-arrested cells, but phosphorylation increased markedly 20 min after release from α-factor, around the time S-CDK is activated and DNA replication begins (Fig. 1d). In a strain in which the two S phase cyclins (*CLB5* and *CLB6*) have been deleted, phosphorylation of T600 after release from α-factor was delayed (to 30 min) and dramatically reduced, indicating that T600 phosphorylation *in vivo* requires S-CDK. Sld3 thus seems to be an essential S-CDK substrate whose phosphorylation is required for DNA replication.

CDK phosphorylation of Sld3 generates a binding site for Dpb11

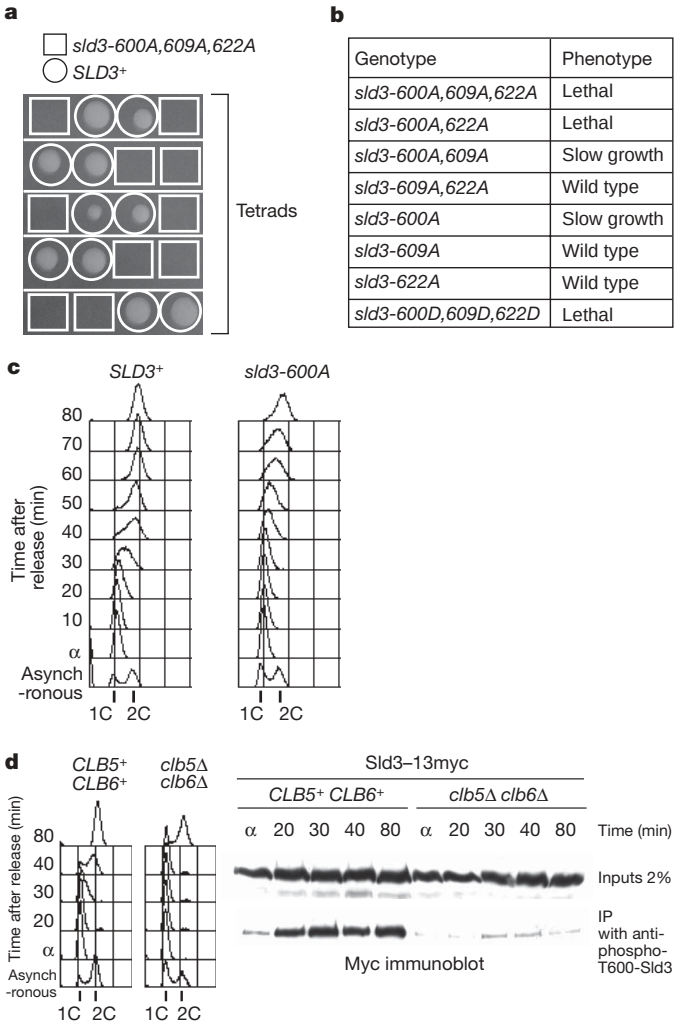
Sld2 and Sld3 were isolated by virtue of genetic interactions with Dpb11 (refs 10, 11, 14) and Dpb11 overexpression can suppress the growth defect of a CDK phosphorylation site mutant of Sld2 (ref. 12). Dpb11 overexpression also suppressed the growth defect (Fig. 2a) and slow S phase (Supplementary Fig. 3) of the *sld3-600A,609A* mutant. These phenotypes were not suppressed by overexpression of Dpb11ΔB1, which lacks amino acid residues 1–95, including one of four BRCT domains of Dpb11 (Fig. 2b).

BRCT domain pairs can bind phosphopeptides[15] and CDK phosphorylation of T84 of Sld2 generates a binding site for the carboxy-terminal two BRCT domains of Dpb11[16]. On the basis of the suppression in Fig. 2a, we predicted that Dpb11 may interact with phosphorylated Sld3. Interactions between recombinant Dpb11 and synthetic Sld3 and Sld2 peptides were measured by surface plasmon resonance. Full-length Dpb11 (1–764) specifically bound to peptides phosphorylated at T600, T609, S622 (Sld3) or T84 (Sld2) (blue and

[1]Cancer Research UK London Research Institute, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK.

©2007 Nature Publishing Group

green lines Fig. 2c, top left panel), but did not bind to unphosphorylated versions of these peptides (red and magenta lines Fig. 2c, top left panel). A C-terminal fragment of Dpb11 spanning BRCT3 and 4 (248–764) bound to the Sld2- but not the Sld3-phosphopeptide (Fig. 2c, top right panel). Conversely, an N-terminal fragment of Dpb11 containing BRCT1 and 2 (1–395) bound specifically to the Sld3- but not the Sld2-phosphopeptide (Fig. 2c, bottom left panel). Mutation of critical residues within either BRCT1 (Fig. 2c, bottom right panel) or BRCT2 (data not shown) of Dpb11 eliminated the specific interaction with the Sld3-phosphopeptide, suggesting that both BRCT1 and 2 are required for this interaction. Figure 2d shows that phospho-T600 or S622 alone have low affinity for Dpb11, but in combination allow more binding to Dpb11. No interactions were observed between Dpb11 and phospho-T609 (data not shown). Peptide pull-downs provided a second approach to show the phosphorylation-dependent interaction of Dpb11 with Sld2 and Sld3 (Fig. 2e). Therefore, there is a correlation between the essential Sld3 residues *in vivo* (T600, S622) and the phospho-dependent interaction with the Dpb11 N-terminus *in vitro*. T609, which has no phenotype when mutated (Fig. 1b), has no role in the interaction with Dpb11 when phosphorylated. Together, our results indicate that phosphorylation of Sld3 at T600 and S622 by S-CDK generates a binding site for the N-terminal BRCT repeats of Dpb11, which is essential for DNA replication.



Genotype	Phenotype
sld3-600A,609A,622A	Lethal
sld3-600A,622A	Lethal
sld3-600A,609A	Slow growth
sld3-609A,622A	Wild type
sld3-600A	Slow growth
sld3-609A	Wild type
sld3-622A	Wild type
sld3-600D,609D,622D	Lethal

Figure 1 | Sld3 is an essential CDK substrate. **a**, Representative tetrads from an SLD3$^+$/*sld3-600A,609A,622A* heterozygote. **b**, Table representing phenotypes from *sld3* mutants generated as in **a**. **c**, Flow cytometry of wild-type *SLD3* yeast (*SLD3*$^+$) or *sld3-600A*, synchronized with α factor and released into nocodazole at 24 °C. DNA contents 1C and 2C refer to the unreplicated and replicated haploid cell DNA contents, respectively. **d**, Myc-tagged-Sld3 strains, either wild type (*CLB5*$^+$,*CLB6*$^+$) or Clb5,6-deleted (*clb5Δ clb6Δ*), were synchronized as in **c** (flow cytometry, left panel). Extracts from this synchronization were used in immunoprecipitation with an antibody specific for phospho-threonine 600 of Sld3. Immunoprecipitates were resolved and detected by Myc-specific western analysis (right panel).

Dpb11 interaction with Sld3 is essential

We predicted that if the primary function of Sld3 phosphorylation is to generate a binding site for Dpb11, we might be able to bypass the essential requirement for the Sld3 CDK sites by enforcing an interaction between Sld3 and Dpb11. To do this, sequence encoding FKBP (FK506 binding protein) was fused to the 3′ end of *sld3-600A,609A,622A* at the endogenous *SLD3* locus (Fig. 3a). As described above (Fig. 1a), this *sld3* mutant is non-functional, so the strain was maintained with an episomal copy of Sld3 on a plasmid containing the *URA3* gene. A second copy of *DPB11* encoded Dpb11 fused to FRB (FRAP 2021–2113) at its N terminus. FKBP interacts with FRB only when bound to the heterobifunctional rapamycin analogue AP21967 (see http://www.ariad.com, for more details). This strain was grown on 5-fluoroorotic acid (5-FOA) plates (to eliminate wild-type *SLD3*) supplemented with AP21967. Surviving colonies were obtained in a strain expressing FRB–Dpb11 but not in a strain expressing Dpb11 without FRB (data not shown). Growth of 5-FOA-selected clones was AP21967-dependent (Fig. 3b). This experiment indicates that enforced interaction between Sld3 and Dpb11 can bypass the lethality of the Sld3 CDK phosphorylation site mutant.

We next tried a more direct approach to bypass the *sld3-600A,609A,622A* mutant. The genomic copy of *sld3-600A,609A,622A* was fused at its 3′-end to sequence encoding a truncated Dpb11

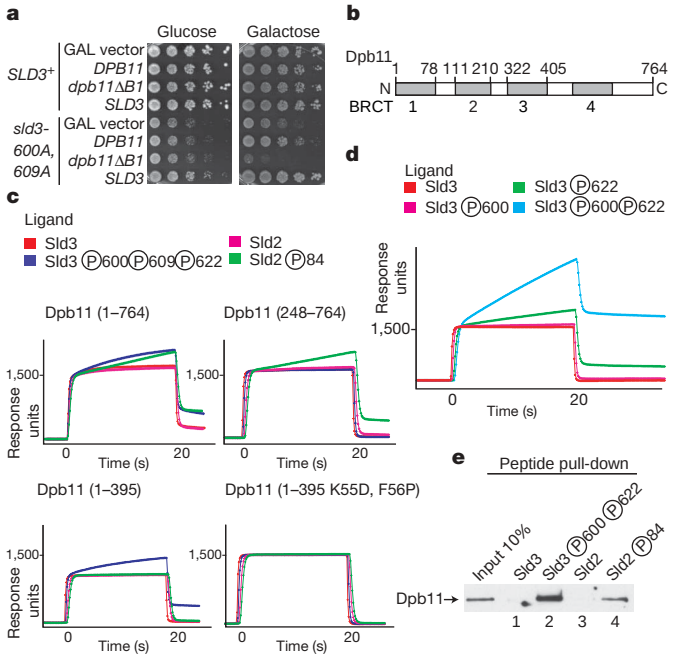


Figure 2 | CDK-phosphorylated Sld3 binds to Dpb11. **a**, Growth assay with *SLD3*$^+$ or *sld3-600A,609A* strains containing the designated galactose-dependent expression constructs. Dpb11ΔB1 is amino acids 96–764 of Dpb11. **b**, Diagram of *S. cerevisiae* Dpb11 showing the position of BRCT repeats. **c**, The four channels of a BIACORE chip were bound with peptide ligand as shown. Sld3 peptide corresponds to amino acids 590–639 and Sld2 peptide to amino acids 61–110. The analytes were recombinant Dpb11–GST fusions as shown. Binding events were measured in response units. **d**, As in **c**, but the analyte was Dpb11(1–764)–GST. **e**, Peptides bound to beads were used to pull down Dpb11(1–764)–GST. Precipitated proteins were analysed by anti-GST western analysis.

©2007 Nature Publishing Group

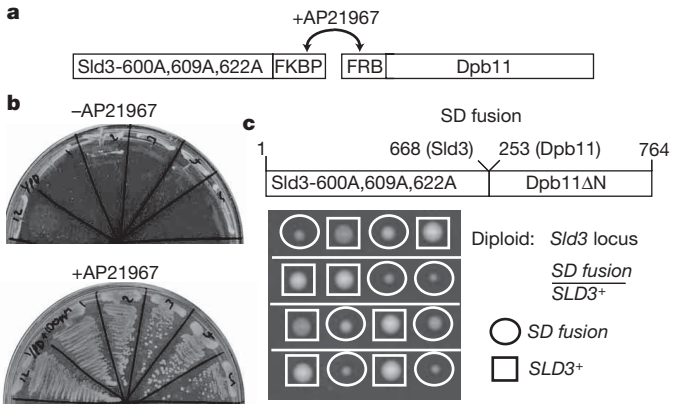
(253–764) lacking BRCT repeats 1 and 2 ('SD fusion'; Fig. 3c). This fragment of Dpb11 cannot bind to phosphorylated Sld3, though it can still bind to CDK-phosphorylated Sld2 (Fig. 2c). Figure 3c shows representative tetrads from an *SLD3*$^+$/*SD fusion* heterozygous diploid and, in contrast with Fig. 1a, all haploid spores from this diploid survive. This experiment confirms that enforced interaction between Sld3 and Dpb11 can bypass the requirement for the otherwise essential Sld3 CDK sites. These results also imply that dissociation of Sld3 and Dpb11 is not essential for viability. Supplementary Fig. 4 shows that the N-terminal BRCT repeats of Dpb11 are essential for viability; however, Fig. 3d shows that *SD fusion*/*dpb11*Δ double mutants are viable. These experiments strongly argue that interaction between Sld3 and Dpb11 is the only essential function of both the CDK phosphorylation of Sld3 and the N-terminal BRCT repeats of Dpb11.

DNA replication without S-CDK

If Sld2 and Sld3 are the only S-CDK substrates essential for DNA replication, bypass of both the phosphorylation of Sld3 and Sld2 should allow for replication in the absence of S-CDK activity. In a haploid strain containing the *SD fusion* as the only copy of *SLD3*, we integrated either *SLD2* wild type or the phospho-mimicking *SLD2-T84D* as a second copy expressed from the inducible galactose promoter. We used expression of a stabilized form of the S- and M-CDK inhibitor Sic1 (Sic1ΔNT)[17–20], after release from an α-factor block, to obtain a synchronous culture of late G1-phase-arrested cells lacking S-CDK. In all strains tested, bud emergence (Fig. 4a)—a G1-CDK-dependent event—occurred with similar kinetics, indicating synchronous release from α-factor arrest. However, Orc6 phosphorylation (Fig. 4b)—an S-CDK-dependent event[21]—did not occur, indicating that S-CDK was efficiently inhibited. No DNA replication occurred in strains with wild-type Sld2 and Sld3, or in strains expressing the Sld2-T84D S-CDK bypass mutant alone. Expression of the SD fusion alone also did not generate substantial DNA replication although small amounts of DNA replication seemed to occur as judged by peak widening and progression towards 2C (Fig. 4c). Remarkably, extensive DNA replication occurred in the presence of both the Sld3- and the Sld2-S-CDK bypass mutants (Fig. 4c). S-CDK-independent DNA replication also occurred when both bypass constructs were expressed from their own promoters (Supplementary Fig. 5), as the only genomic copy. S-CDK-independent replication continued well beyond 2C, suggesting considerable re-replication. This is not unexpected because low CDK activity should allow origin resetting after firing[2]. Supplementary Fig. 6 shows that this S-CDK-independent DNA replication still requires Cdc7 function, indicating that Cdc7 functions independently of S-CDK. Therefore, bypassing the requirement for CDK phosphorylation of Sld2 and Sld3 is sufficient to bypass the requirement for S-CDK in DNA replication.

Additional requirement for DNA replication in G1 phase

The loss of all three G1 cyclins is lethal[22] but is suppressed by deletion of Sic1 (refs 23, 24), suggesting that the essential role of G1-CDK is to promote activation of S-CDK by eliminating Sic1 (refs 25–27). If true, bypassing the requirement for S-CDK in DNA replication as described above should also bypass the requirement for G1-CDK. Figure 5a shows that, although expression of the Sld2 and Sld3 bypass mutants induces a small amount of DNA synthesis in α-factor-arrested cells (right panel), it is much less than the DNA synthesis induced in Sic1ΔNT-arrested cells (left panel), strongly suggesting that G1-CDK is required for something in addition to activation of S-CDK.



d

Genotype	Spores (%)
SLD3/DPB11	30
*SLD3/dpb11*Δ	0
SD fusion/DPB11	38
*SD fusion/dpb11*Δ	32

Diploid: *Sld3* locus *SD fusion*/*SLD3*$^+$ *Dpb11* locus *dpb11*Δ/*DPB11*$^+$

Figure 3 | Dpb11–Sld3 interaction is the essential function of Sld3 CDK phosphorylation sites and the N terminus of Dpb11. a, Diagram of strain construction for **b**. FRB binds to FKBP in the presence of drug AP21967. The strain is maintained by an episomal *URA3* plasmid expressing wild-type Sld3. **b**, Colonies were selected on 5-FOA with AP21967, and restreaked with or without AP21967. **c**, Schematic of fusion between sld3-600A,609A,622A and the C terminus of Dpb11 (top panel), and tetrads from an *SD fusion*/*SLD3*$^+$ heterozygote. **d**, Genotypes of haploids from an *SD fusion* and Dpb11 deletion (*dpb11*Δ) heterozygote ($n = 72$).

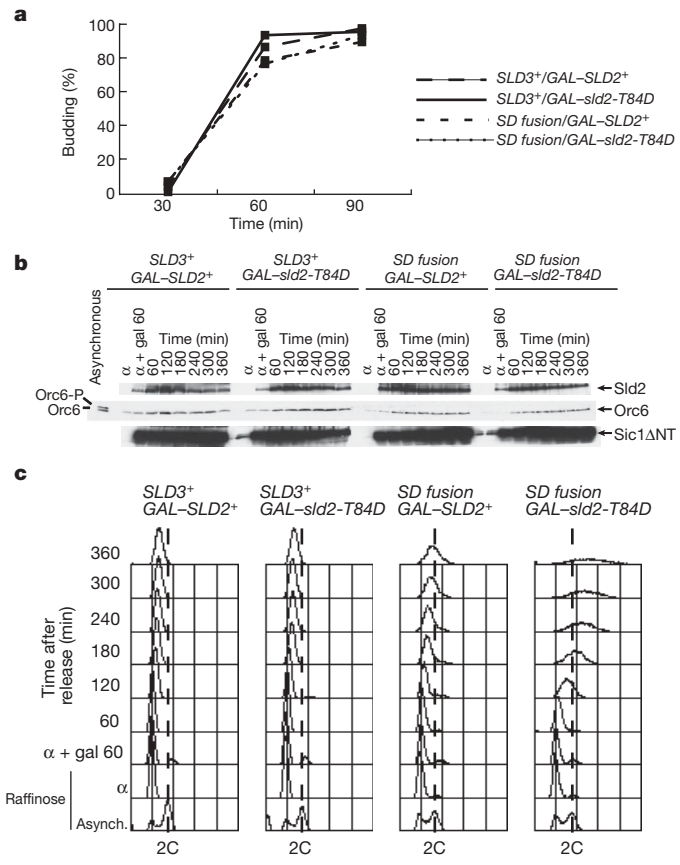
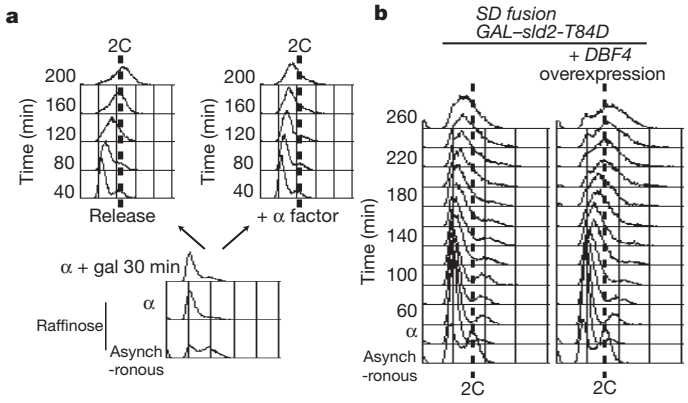
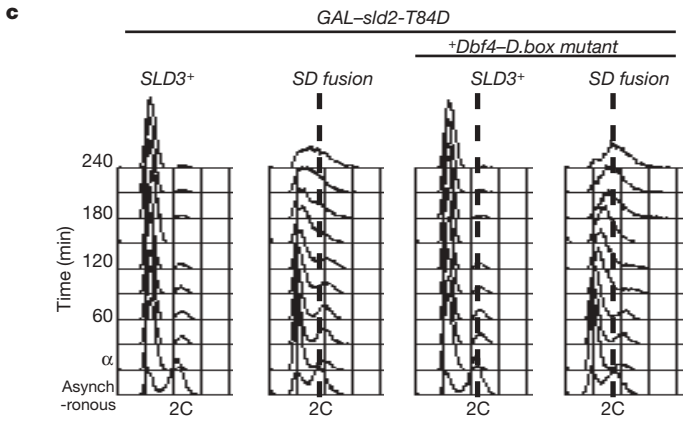


Figure 4 | DNA replication in the absence of S-CDK activity. a, Strains with wild-type Sld3 (*SLD3*$^+$) or *SD fusion* with Sld2 (either wild type or T84D) expressed from a galactose dependent promoter were synchronized in raffinose with α factor and released into S phase in galactose, and budding was recorded. All strains express Sic1ΔNT from a galactose dependent promoter. **b**, Immunoblotting of protein extracts from experiment in **a**. Orc6-P is the S-CDK phosphorylated form of Orc6. **c**, Flow cytometry of cells from the experiment in **a**.

©2007 Nature Publishing Group





d

Genotype	Spores (%) $n = 72$
SLD3⁺/MCM5⁺	27.8
SLD3⁺/mcm5-bob1	44.4
SD fusion/MCM5⁺	27.8
SD fusion/mcm5-bob1	0

Sld3 locus: *SD fusion* / *SLD3⁺*
Mcm5 locus: *mcm5-bob1* / *MCM5⁺*
Cdc7 locus: *cdc7Δ* / *CDC7⁺*

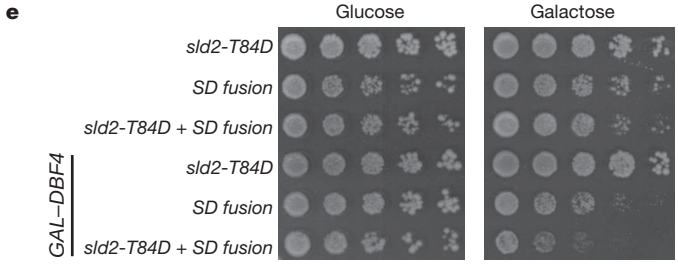


Figure 5 | CDK bypass and Dbf4 overexpression allows replication in G1 cells. a, The CDK bypass strain containing the *SD fusion*, *GAL–sld2-T84D* and *GAL–Sic1ΔNT* was synchronized with α factor (lower panel) and either released into S phase or held in α factor in galactose. **b**, The strain from **a** (with *BAR1* deleted), was synchronized with and held in α factor in galactose with or without Dbf4 overexpression from a galactose-dependent promoter. **c**, As in **b** except that the strains in the two right panels express Dbf4 with both APC/C destruction boxes mutated (*dbf4-D.box mutant*) from the *DBF4* promoter. **d**, Genotypes of haploids from an *SD fusion*, *mcm5-bob1* and *cdc7Δ* heterozygote ($n = 72$). **e**, Growth assay with strains containing *sld2-T84D* and/or *SD fusion* as the only genomic copy expressed from their own promoter, with or without Dbf4 expressed from a galactose-dependent promoter.

G1-CDK promotes activation of S-CDK not only by targeting Sic1 for degradation[25–27] but also by inactivating the Anaphase Promoting Complex/Cyclosome (APC/C)[28]—the E3 ubiquitin ligase required for cyclin degradation. In addition to targeting the S and M cyclins, APC/C targets the regulatory subunit of the Cdc7 kinase—Dbf4—for proteolysis, and Dbf4 is present at very low levels in α-factor-arrested cells[29–31]. Figure 5b shows that galactose-dependent overexpression of Dbf4 together with the Sld2 and Sld3 S-CDK-bypass mutants promotes more extensive DNA synthesis in α-factor-arrested cells. Figure 5c shows that the stabilized destruction box mutant of Dbf4 (ref. 31), expressed from the *DBF4* promoter, also promotes more DNA replication in α-factor-arrested cells. Therefore, Dbf4 is limiting for DNA replication in α-factor-arrested cells when the requirement for S-CDK is bypassed.

To assess the importance of this, we asked whether combining S-CDK bypass with Cdc7 bypass affects viability. Figure 5d shows that the SD fusion, which by itself can weakly bypass S-CDK requirement (Fig. 4c), is synthetically lethal with *mcm5-bob1*, a mutant that bypasses the requirement for Cdc7 (ref. 32). Synthetic lethality was independent of Cdc7 status, indicating that it is the bypass of Cdc7 requirement rather than the absence of Cdc7 that is lethal. Next we examined the effect of Dbf4 deregulation in a strain containing the *SD fusion* and *SLD2-T84D* expressed from their own promoters. Figure 5e shows that *SLD2-T84D* grows normally whereas the *SD fusion* alone has a slight growth defect (see also Fig. 3c). Interestingly, the *SLD2-T84D* strain with the *SD fusion* together is also viable and grows similarly to the *SD fusion* alone, indicating that bypass of S-CDK requirement is not lethal by itself. Overexpression of Dbf4 had no affect on *SLD2-T84D*, slightly reduced growth of the *SD fusion* strain and greatly reduced growth of the double-bypass strain (Fig. 5e). Finally, constructs expressing the Dbf4 destruction box mutant from its own promoter efficiently transformed either of the single S-CDK bypass mutants, but the same construct transformed the double-S-CDK-bypass strain much less efficiently (Supplementary Fig. 7). Together, these data argue that regulation of Dbf4/Cdc7 and S-CDK act redundantly to prevent premature DNA replication in G1 phase.

Discussion

Using a directed approach, we have shown that Sld2 and Sld3 are the minimal set of essential S-CDK substrates required for DNA replication in budding yeast. S-CDK phosphorylation of additional proteins may contribute to efficient DNA replication or may be required for the myriad of replication-coupled processes such as sister chromatid cohesion, DNA damage signalling, post-replication repair, chromatin assembly and so on. Our system for inducing DNA replication without CDKs should provide an approach to examine this. Further detailed biochemistry will be required to determine how the phosphorylation-dependent binding of Dpb11 to Sld2 and Sld3 promotes subsequent origin unwinding and replisome assembly.

The ability of the Sld2 and Sld3 bypass mutants together with Dbf4 to induce DNA replication in α-factor-arrested cells completes a pathway connecting an extracellular growth factor (α factor) to the initiation of DNA replication. This system should allow us to assess the contribution of other G1-CDK-regulated functions such as late G1 phase transcription. Our results indicate that the downregulation of S-CDK and Dbf4/Cdc7 both contribute to preventing inappropriate DNA replication during G1 phase. Premature replication in G1 phase may be lethal because it uncouples DNA replication from growth control, because it uncouples replication from the block to re-replication or because it uncouples DNA replication from other S-CDK-dependent events.

To date, Sld3 homologues have not been identified in multicellular eukaryotes, though a putative Sld2 orthologue has been found[33,34]. Indeed, even within fungi, Sld3 is very poorly conserved, suggesting rapid evolution. Because the N-terminal BRCT repeats of Dpb11 are conserved in its metazoan orthologues (TopBP1 or Mus101)[35,36] it is

©2007 Nature Publishing Group

likely that an Sld3 orthologue will be found in metazoans. Further work is required to determine if S-CDK targets are also conserved in evolution.

METHODS

See Supplementary Information for experimental details.

Yeast strains and methods. All yeast strains were based on W303-1a (*MAT**a** ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100*) and are described in the Supplementary Table.

Cell growth and cell cycle blocks, samples for flow-cytometric analysis and yeast protein extracts were as previously described[37–39].

Drug-dependent heterodimerization of Sld3 and Dpb11. Reagents were derived from the ARGENT Regulated Heterodimerization Kit (ARIAD pharmaceuticals, http://www.ariad.com/). FKBP (FK506 binding protein) and FRB (FRAP 2021–2113 T2098L) tags were subcloned from supplied plasmids and rapamycin analogue AP21967 was used at 100 nM. The 5-FOA concentration was 1 mg ml^{-1}.

Surface plasmon resonance (SPR). SPR was done using a BIACORE 3000. Peptides were synthesized with an N-terminal biotin moiety and were coupled directly to streptavidin-coated BIACORE sensor chips.

Phospho-antibody affinity purification and immunoprecipitation. Rabbit polyclonal antibodies, raised against Sld3 peptide (596–604) phosphorylated at threonine 600, were affinity purified with the antigen peptide as described[40]. Extracts were made from trichloroacetic acid (TCA)-treated yeast cells by boiling in 0.5% SDS, 10% glycerol, 50 mM Tris-HCl pH 6.8 and then diluting to 0.1% SDS, 150 mM NaCl, 1% NP40, 0.5% DOC, 50 mM Tris pH 8. Anti-phospho-Sld3 antibody was incubated with extract overnight and precipitated proteins were resolved by SDS polyacrylamide gel electrophoresis.

Received 31 July; accepted 10 November 2006.
Published online 13 December 2006.

1. Blow, J. J. & Dutta, A. Preventing re-replication of chromosomal DNA. *Nature Rev. Mol. Cell Biol.* **6,** 476–486 (2005).
2. Diffley, J. F. X. Regulation of early events in chromosome replication. *Curr. Biol.* **14,** R778–R786 (2004).
3. Donaldson, A. D. *et al.* CLB5-dependent activation of late replication origins in *S. cerevisiae*. *Mol. Cell* **2,** 173–183 (1998).
4. Epstein, C. B. & Cross, F. R. *CLB5*, a novel B cyclin from budding yeast with a role in S phase. *Genes Dev.* **6,** 1695–1706 (1992).
5. Kuhne, C. & Linder, P. A new pair of B-type cyclins from *Saccharomyces cerevisiae* that function early in the cell cycle. *EMBO J.* **12,** 3437–3447 (1993).
6. Schwob, E. & Nasmyth, K. *CLB5* and *CLB6*. a new pair of B cyclins involved in DNA replication in *Saccharomyces cerevisiae*. *Genes Dev.* **7,** 1160–1175 (1993).
7. Bousset, K. & Diffley, J. F. The Cdc7 protein kinase is required for origin firing during S phase. *Genes Dev.* **12,** 480–490 (1998).
8. Donaldson, A. D., Fangman, W. L. & Brewer, B. J. Cdc7 is required throughout the yeast S phase to activate replication origins. *Genes Dev.* **12,** 491–501 (1998).
9. Ubersax, J. A. *et al.* Targets of the cyclin-dependent kinase Cdk1. *Nature* **425,** 859–864 (2003).
10. Wang, H. & Elledge, S. J. DRC1, DNA replication and checkpoint protein 1, functions with DPB11 to control DNA replication and the S-phase checkpoint in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **96,** 3824–3829 (1999).
11. Kamimura, Y., Masumoto, H., Sugino, A. & Araki, H. Sld2, which interacts with Dpb11 in *Saccharomyces cerevisiae*, is required for chromosomal DNA replication. *Mol. Cell. Biol.* **18,** 6102–6109 (1998).
12. Masumoto, H., Muramatsu, S., Kamimura, Y. & Araki, H. S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature* **415,** 651–655 (2002).
13. Noguchi, E., Shanahan, P., Noguchi, C. & Russell, P. CDK phosphorylation of Drc1 regulates DNA replication in fission yeast. *Curr. Biol.* **12,** 599–605 (2002).
14. Kamimura, Y., Tak, Y. S., Sugino, A. & Araki, H. Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in *Saccharomyces cerevisiae*. *EMBO J.* **20,** 2097–2107 (2001).
15. Glover, J. N., Williams, R. S. & Lee, M. S. Interactions between BRCT repeats and phosphoproteins: tangled up in two. *Trends Biochem. Sci.* **29,** 579–585 (2004).
16. Tak, Y. S., Tanaka, Y., Endo, S., Kamimura, Y. & Araki, H. A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2–Dpb11. *EMBO J.* **25,** 1987–1996 (2006).
17. Schwob, E., Bohm, T., Mendenhall, M. D. & Nasmyth, K. The B-type cyclin kinase inhibitor p40^{SIC1} controls the G1 to S transition in *S. cerevisiae*. *Cell* **79,** 233–244 (1994).
18. Desdouets, C. *et al.* Evidence for a Cdc6p-independent mitotic resetting event involving DNA polymerase α. *EMBO J.* **17,** 4139–4146 (1998).
19. Noton, E. A. & Diffley, J. F. X. CDK inactivation is the only essential function of the APC/C and the mitotic exit network proteins for origin resetting during mitosis. *Mol. Cell* **5,** 85–95 (2000).
20. Tanaka, S. & Diffley, J. F. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2–7 during G1 phase. *Nature Cell Biol.* **4,** 198–207 (2002).
21. Wilmes, G. M. *et al.* Interaction of the S-phase cyclin Clb5 with an 'RXL' docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes Dev.* **18,** 981–991 (2004).
22. Richardson, H., Wittenberg, C., Cross, F. & Reed, S. An essential G1 function for cyclin-like proteins in yeast. *Cell* **59,** 1127–1133 (1989).
23. Tyers, M. The cyclin-dependent kinase inhibitor p40^{SIC1} imposes the requirement for Cln G1 cyclin function at Start. *Proc. Natl Acad. Sci. USA* **93,** 7772–7776 (1996).
24. Schneider, B. L., Yang, Q. H. & Futcher, A. B. Linkage of replication to start by the Cdk inhibitor Sic1. *Science* **272,** 560–562 (1996).
25. Nash, P. *et al.* Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* **414,** 514–521 (2001).
26. Patton, E. E. *et al.* Cdc53 is a scaffold protein for multiple Cdc34/Skp1/F-box protein complexes that regulate cell division and methionine biosynthesis in yeast. *Genes Dev.* **12,** 692–705 (1998).
27. Verma, R. *et al.* Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S Phase. *Science* **278,** 455–460 (1997).
28. Zachariae, W., Schwab, M., Nasmyth, K. & Seufert, W. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science* **282,** 1721–1724 (1998).
29. Cheng, L., Collyer, T. & Hardy, C. F. Cell cycle regulation of DNA replication initiator factor Dbf4p. *Mol. Cell. Biol.* **19,** 4270–4278 (1999).
30. Oshiro, G., Owens, J. C., Shellman, Y., Sclafani, R. A. & Li, J. J. Cell cycle control of Cdc7p kinase activity through regulation of Dbf4p stability. *Mol. Cell. Biol.* **19,** 4888–4896 (1999).
31. Godinho Ferreira, M., Santocanale, C., Drury, L. S. & Diffley, J. F. X. Dbf4p, an essential S phase promoting factor, is targeted for degradation by the Anaphase Promoting Complex. *Mol. Cell. Biol.* **20,** 242–248 (2000).
32. Hardy, C. F., Dryga, O., Seematter, S., Pahl, P. M. & Sclafani, R. A. Mcm5/Cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. *Proc. Natl Acad. Sci. USA* **94,** 3151–3155 (1997).
33. Matsuno, K., Kumano, M., Kubota, Y., Hashimoto, Y. & Takisawa, H. The N-terminal noncatalytic region of *Xenopus* RecQ4 is required for chromatin binding of DNA polymerase α in the initiation of DNA replication. *Mol. Cell. Biol.* **26,** 4843–4852 (2006).
34. Sangrithi, M. N. *et al.* Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell* **121,** 887–898 (2005).
35. Makiniemi, M. *et al.* BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *J. Biol. Chem.* **276,** 30399–30406 (2001).
36. Yamamoto, R. R. *et al.* The Drosophila *mus101* gene, which links DNA repair, replication and condensation of heterochromatin in mitosis, encodes a protein with seven BRCA1 C-terminus domains. *Genetics* **156,** 711–721 (2000).
37. Diffley, J. F., Cocker, J. H., Dowell, S. J. & Rowley, A. Two steps in the assembly of complexes at yeast replication origins *in vivo*. *Cell* **78,** 303–316 (1994).
38. Labib, K., Diffley, J. F. & Kearsey, S. E. G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nature Cell Biol.* **1,** 415–422 (1999).
39. Foiani, M., Marini, F., Gamba, D., Lucchini, G. & Plevani, P. The B subunit of the DNA polymerase alpha-primase complex in *Saccharomyces cerevisiae* executes an essential function at the initial stage of DNA replication. *Mol. Cell. Biol.* **14,** 923–933 (1994).
40. Harlow, E. & Lane, D. *Using Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. Araki and colleagues for sharing unpublished information. We are grateful to J. A. Tercero for making strain y1187 and S. Sweet for making strain y2003 and providing extracts. We are grateful to ARIAD pharmaceuticals for providing the ARGENT Regulated Heterodimerization Kit. We thank J. Gannon for assistance with the BIACORE 3000. We are grateful to N. O'Reilly and the Peptide Synthesis Facility at the London Research Institute. We thank members of our laboratory for helpful discussion and criticism of the manuscript. P.Z. is supported by a Cancer Research UK fellowship.

Author Contributions P.Z. and J.F.X.D. conceived and designed experiments and wrote the paper. P.Z. performed all experiments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.F.X.D. (John.Diffley@cancer.org.uk).

©2007 Nature Publishing Group

LETTERS

Dark matter maps reveal cosmic scaffolding

Richard Massey[1], Jason Rhodes[1,2], Richard Ellis[1], Nick Scoville[1], Alexie Leauthaud[3], Alexis Finoguenov[4], Peter Capak[1], David Bacon[5], Hervé Aussel[6], Jean-Paul Kneib[3], Anton Koekemoer[7], Henry McCracken[8], Bahram Mobasher[7], Sandrine Pires[9], Alexandre Refregier[6], Shunji Sasaki[10], Jean-Luc Starck[9], Yoshi Taniguchi[10], Andy Taylor[5] & James Taylor[11]

Ordinary baryonic particles (such as protons and neutrons) account for only one-sixth of the total matter in the Universe[1–3]. The remainder is a mysterious 'dark matter' component, which does not interact via electromagnetism and thus neither emits nor reflects light. As dark matter cannot be seen directly using traditional observations, very little is currently known about its properties. It does interact via gravity, and is most effectively probed through gravitational lensing: the deflection of light from distant galaxies by the gravitational attraction of foreground mass concentrations[4,5]. This is a purely geometrical effect that is free of astrophysical assumptions and sensitive to all matter—whether baryonic or dark[6,7]. Here we show high-fidelity maps of the large-scale distribution of dark matter, resolved in both angle and depth. We find a loose network of filaments, growing over time, which intersect in massive structures at the locations of clusters of galaxies. Our results are consistent with predictions of gravitationally induced structure formation[8,9], in which the initial, smooth distribution of dark matter collapses into filaments then into clusters, forming a gravitational scaffold into which gas can accumulate, and stars can be built[10].

The Hubble Space Telescope (HST) Cosmic Evolution Survey (COSMOS) is the largest contiguous expanse of high-resolution astronomical imaging obtained from space[11]. 575 slightly overlapping pointings of the Advanced Camera for Surveys (ACS) Wide Field Camera cover a region of 1.637 square degrees (more details about the COSMOS survey are available at www.astro.caltech.edu/~cosmos). We measure the shapes of half a million distant galaxies[12], and use their observed distortion (compare with Supplementary Fig. 1) to reconstruct the distribution of intervening mass, projected along our line of sight (Fig. 1). A realization of noise in our mass map, including most spurious instrumental or systematic effects, is provided by the '*B*-mode' signal. This is an additional degree of freedom in the data, which is not produced by gravitational lensing, and so is expected to be zero in the absence of systematics[13]. Assuming a gaussian noise distribution, the relative number of pixels above and below the first contour in the *B*-mode suggests that this

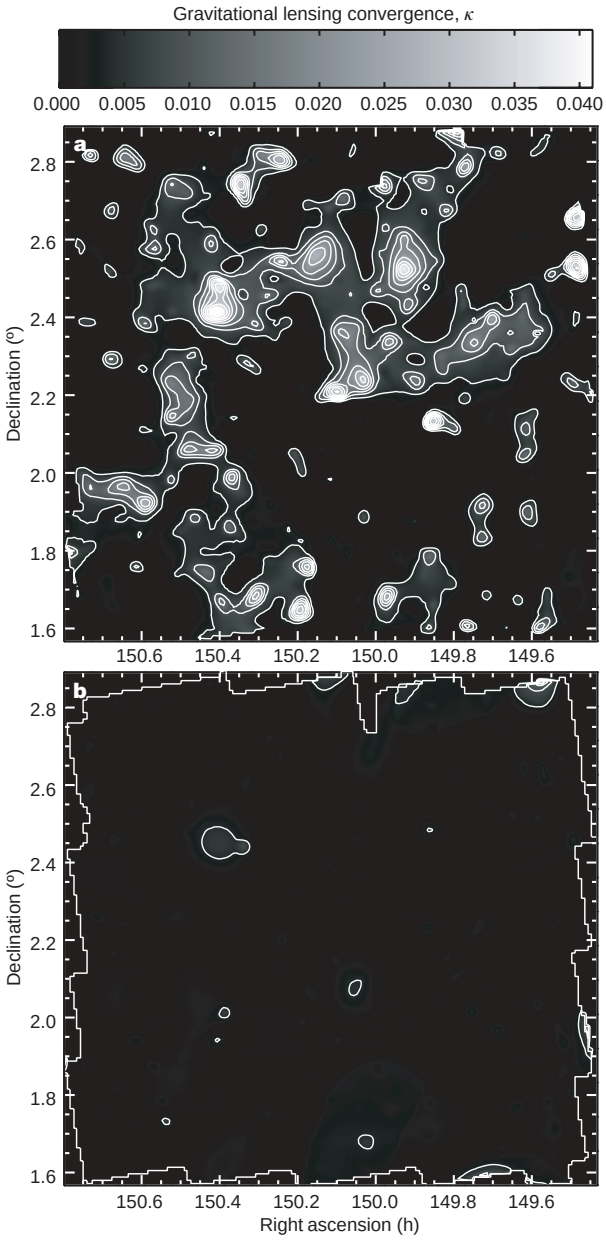


Figure 1 | Map of the dark matter distribution in the two-square-degrees COSMOS field. a, The linear greyscale shows the *E*-mode lensing convergence field κ, which is proportional to the projected mass along the line of sight. Contours begin at 0.4% and are spaced by 0.5% in κ. **b**, The absolute value of the *B*-mode signal, shown with the same greyscale and contour levels, provides a realization of the noise level in the map, plus contamination from uncorrected systematic effects; the bold outline traces the region observed with HST.

[1]California Institute of Technology MC105-24, 1200 E. California Boulevard, Pasadena, California 91125, USA. [2]Jet Propulsion Laboratory, Pasadena, California 91109, USA. [3]Laboratoire d'Astrophysique de Marseille, 13376 Marseille Cedex 12, France. [4]Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße, 85748 Garching, Germany. [5]Institute for Astronomy, Blackford Hill, Edinburgh EH9 3HJ, UK. [6]AIM, Unité Mixte de Recherche CEA, CNRS et Université de Paris VII, UMR no. 7158 CE Saclay, 91191 Gif-sur-Yvette, France. [7]Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. [8]Institut d'Astrophysique de Paris, Université Pierre et Marie Curie, 98 bis Boulevard Arago, 75014 Paris, France. [9]CEA/DSM/DAPNIA/SEDI, CE Saclay, 91191 Gif-sur-Yvette, France. [10]Physics Department, Ehime University, 2-5 Bunkyou, Matuyama 790-8577, Japan. [11]Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

©2007 Nature Publishing Group

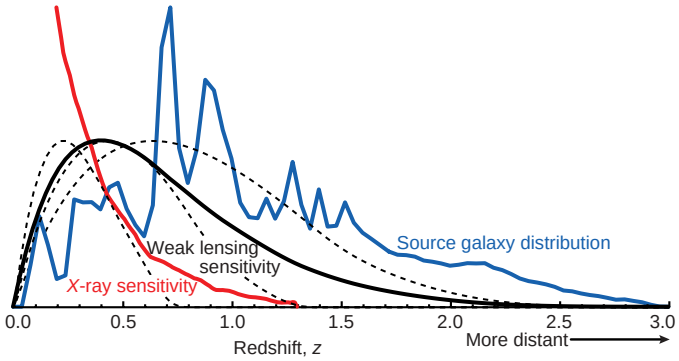


Figure 2 | Sensitivity of probes of large-scale structure, as a function of distance. The blue line shows the distribution of photometric redshifts for the source galaxies. The solid black line shows the sensitivity function of the lensing material for this source population (arbitrarily normalized to peak at unity), whereas the dashed lines show the equivalent sensitivities for the tomographic analysis. The red line shows the (arbitrarily normalized) sensitivity of the X-ray detections. Because the survey volume is a cone, the effective volumes peak at $z = 0.7$ (lensing) and $z = 0.4$ (X-rays).

contour (in both E-mode and B-mode maps) is equivalent to a $\sim 2.9\sigma$ detection threshold that includes both statistical and systematic noise.

Gravitational lensing measurements have an unusual sensitivity to the distance of the influencing mass (Fig. 2), in contrast to the familiar $\propto 1/d^2$ decline in luminosity of optically visible sources. Like an ordinary glass lens, a gravitational lens is most efficient when placed

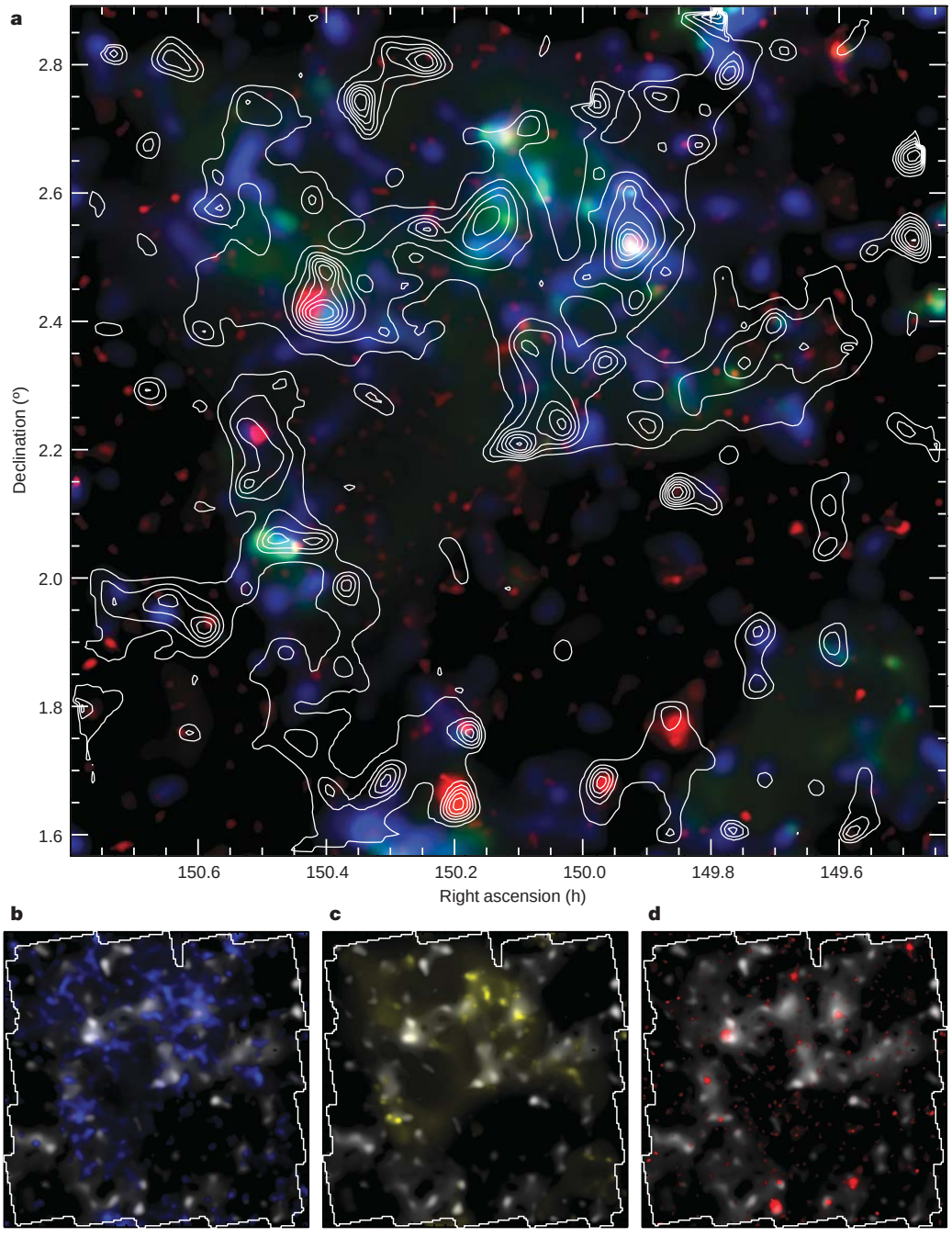


Figure 3 | Comparison of baryonic and non-baryonic large-scale structure. The total projected mass from weak lensing, dominated by dark matter, is shown as contours in **a** and as a linear greyscale in **b**, **c** and **d**. Independent baryonic tracers comprise (1) stellar mass (blue, the colour scale peaks at $2.3 \times 10^{11} M_{\rm Sun}\,{\rm deg}^{-2}$ within $\Delta z = 0.1$), (2) galaxy number density (yellow, peak at $1.4 \times 10^{5}\,{\rm deg}^{-2}$ within $\Delta z = 0.1$) seen in optical and near-infrared light (adjusted to the redshift sensitivity function of the lensing mass), and (3) hot gas (red, peak at $2.6 \times 10^{-14}\,{\rm erg\,s^{-1}\,cm^{-2}\,arcmin^{-2}}$) seen in X-rays after removal of point sources.

©2007 Nature Publishing Group

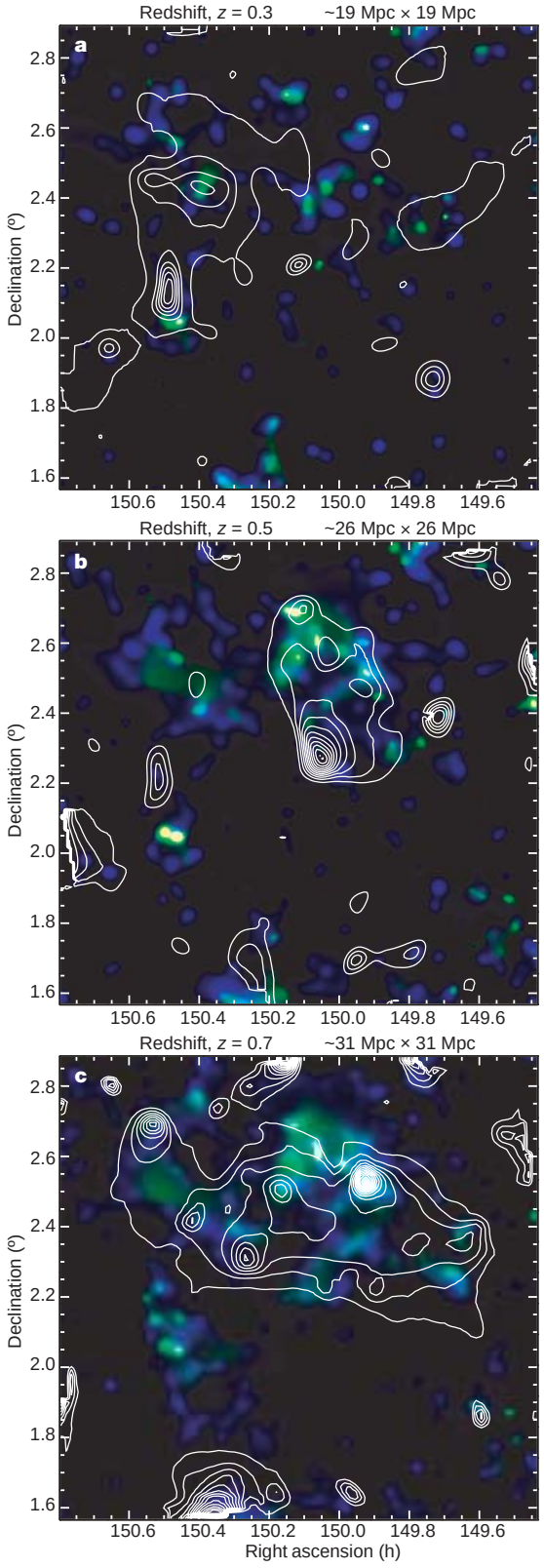


Figure 4 | Growth of large-scale structure. Slices through the evolving distribution of dark matter, created by splitting the background source galaxy population into discrete redshift slices. The sensitivity functions of the mass reconstruction peak at redshifts of ~0.3, ~0.5 and ~0.7 from **a** to **c**. Contours show the lensing convergence, in steps of 0.33%. A linear green colour ramp shows the distribution of galaxies, and blue shows their stellar mass; both are weighted with matched redshift sensitivity functions.

halfway between the source and the observer—so lensing is sensitive to neither very distant, nor very nearby structures. To compare the distribution of baryons with that of dark matter, we mimic this effect by appropriately weighting the foreground galaxies as a function of their redshift (cosmological distance). The number density of these independent galaxies, and the mass contained in their stars (estimated from their colour) provide two matched tracers of baryons. Deep X-ray observations with the XMM-Newton satellite[14] additionally highlight concentrations of hot, dense gas.

The most prominent peak in the projected, two-dimensional distributions of all four tracers (Fig. 3) is a single cluster of galaxies at (149 h 55 min, 2° 31′) and redshift $z = 0.73$. X-rays are sensitive to the square of the electron density, so they preferentially highlight the central cluster core. This cluster has an X-ray temperature $kT_X = 3.51^{+0.60}_{-0.46}$ keV and luminosity $L_X = (1.56 \pm 0.04) \times 10^{44}$ erg s^{-1} (0.1–2.4 keV band)[15]. If the cluster gas distribution were in hydrostatic equilibrium ('relaxed'), this would imply a mass of $(1.6 \pm 0.4) \times 10^{14} M_{Sun}$ within an r_{500} radius of 1.4 arcminutes. However, the cluster is clearly still growing[15]. Gravitational lensing is linearly sensitive to mass, and reveals an extended dark matter halo around this cluster, which in turn lies at the nexus of several filaments. The lensing mass of the full halo is $(6 \pm 3) \times 10^{15} M_{Sun}$. It is possible that such a large value includes a contribution from additional mass directly in front of the cluster, at redshifts where lensing is more sensitive. Similar projection effects might also explain the twin lensing peaks without obvious baryonic counterparts near (150 h 20 min, 2° 40′). Weak lensing analysis is very sensitive, and the map could also have been perturbed by finite-field edge effects or isolated defects in our model of the telescope's point spread function that are difficult to detect individually.

A statistical comparison across the entire map shows that baryons follow the distribution of dark matter even on large scales. The linear regression correlation coefficient r of lensing mass with stellar mass is 0.42 and r of lensing mass with galaxy number density is 0.47. The

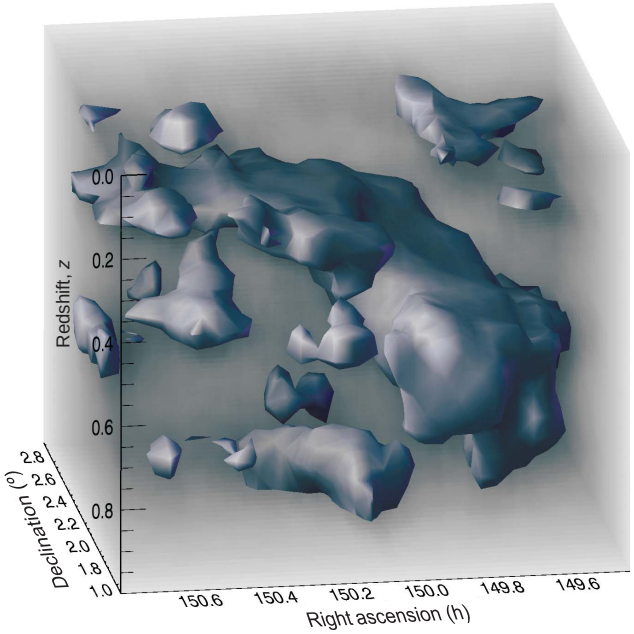


Figure 5 | Three-dimensional reconstruction of the dark matter distribution. The three axes correspond to right ascension, declination and redshift: with distance from the Earth increasing towards the bottom. The redshift scale is highly compressed, and the survey volume is really an elongated cone. An isodensity contour has been drawn at a level of $1.4 \times 10^{13} M_{Sun}$ within a circle of radius 700 kpc and $\Delta z = 0.05$. This was chosen arbitrarily to highlight the filamentary structure. The faint background shows the full distribution, with the level of the greyscale corresponding to the local density. Additional views are provided in Supplementary Fig. 7.

©2007 Nature Publishing Group

correlation with X-ray flux is somewhat lower, 0.30, consistent with the presence of filamentary structure outside the cluster cores. The map reveals overdense regions that are topologically connected but insufficiently dense to generate X-ray emission. These filaments are not a smoothing artefact, and cannot be reproduced by adding noise to the (square root of the) X-ray image and then smoothing in a comparable way. We identify three distinct sets of environments, with a stark mass contrast. Filaments, defined as regions outside clusters but with a lensing magnification $\kappa > 0.4\%$, contain a projected number density of galaxies twice as high as that in voids, and 1.5 times lower than that in X-ray luminous clusters. Filaments have gravitationally collapsed along two axes, but clusters have continued to collapse along the third, and the galaxy number density in clusters is expected to be about five times that of filaments in the plane of the sky[16]. The observed ratio is lowered by noise in the mass reconstruction, as well as partial alignment of filaments along our line of sight.

The three-dimensional distribution of dark matter, and hence its time-dependent growth, can be visualized (Fig. 4) by splitting the background source galaxies into discrete redshift bins[17]. We have chosen bins so that the resulting foreground lensing sensitivities peak at redshifts $z \approx 0.3$, 0.5 and 0.7 (Fig. 2). These functions overlap slightly, so some structures can faintly be seen in successive slices. Catastrophic failures in photometric redshift measurement potentially mix slices further, although we have developed a method that minimizes this effect in the lensing analysis. The massive $z = 0.73$ cluster is indeed part of a much larger three-dimensional structure, including a filament partially aligned with our line of sight, which will increase its two-dimensional projected mass but not affect the X-ray flux. The corresponding *B*-mode maps (Supplementary Fig. 4) suggest that the second contours here have roughly the same 3σ significance as the first contour in Fig. 1. A full three-dimensional reconstruction of the mass distribution (Fig. 5) is obtained from the differential growth of the lensing signal between many thin slices separated by $\Delta z = 0.05$ (refs 18, 19). The evolution of this distribution is driven by the battle between gravitational collapse and the accelerating expansion of the Universe.

The independent probes of large-scale structure paint a remarkably consistent picture of the Universe on large scales. The contracting filamentary network resembles predictions from *n*-body simulations of structure formation dominated by the gravitational collapse of cold dark matter from small density perturbations in the early Universe[9]. By directly probing the distribution of mass, space-based weak lensing measurements offer the potential to link observations directly to theories that are concerned mainly with collisionless dark matter and gravity. Indeed, the resemblance of Fig. 5 to the first three-dimensional maps of the large-scale distribution of baryonic matter made by the Point Source Redshift Catalogue[20] (PSCz) fifteen years ago demonstrates substantial progress in observational astronomy.

METHODS

Shear measurement. The depth and exquisite resolution of Hubble Space Telescope images enable us to resolve the shapes of 71 galaxies per square arcminute in the F814W (approximately *I*-band) filter, with a median AB magnitude of 25.1. We use the RRG method[21] to deconvolve the galaxy shapes from the telescope's point spread function. Our processing pipeline has been calibrated on simulated HST images, and found to recover shear from galaxies in a wide range of size and flux, with less than 6% bias[12].

Mass reconstruction. The observed shear field is then converted into 'convergence' (the cumulative magnification of all lenses along a line of sight), which is proportional to the two-dimensional, projected mass[22]. The conversion is non-local, and finite-field effects introduce some defects near the edge of the map. Diversity in galaxies' intrinsic morphologies propagates into shot noise in the mass map, which we reduce by using a multiscale filtering method, based on the *à trous* wavelet transform[23], and tuned to balance completeness with minimal false detection of spurious signals[24]. Our high surface density of resolved galaxies permits a mass reconstruction with an unprecedented minimum wavelet scale (maximum resolution) of 1.2 arcminutes full-width at half-maximum (FWHM) in the projected map, and 2.4 arcminutes FWHM in the tomographic analysis. In practice, the achieved resolution of the wavelet reconstruction varies spatially with the local signal strength. Noise properties vary on the different wavelet scales, and are summed in a complex fashion that also depends upon the local signal. The noise is best quantified via the corresponding *B*-mode maps.

Charge transfer efficiency correction. One particularly troublesome systematic is introduced by image trailing during charge-coupled device (CCD) detector readout, owing to radiation damage that degrades their charge transfer efficiency (CTE)[25]. The arrangement of the ACS CCDs produces a spurious sawtooth *E*-mode convergence pattern (with an average peak signal of $\pm 0.3\%$ and a pitch of 3.3 arcminutes, corresponding to the ACS field of view); but no corresponding *B*-mode (Supplementary Fig. 2). We subtract a model[26] of the spurious signal, which incorporates parameters of galaxy flux, position on the CCD and date of exposure, from the galaxy shear catalogue. After this correction, and the removal of high-frequency information by wavelet filtering, the spurious *E*-mode convergence signal is less than 0.1% throughout the map (Supplementary Fig. 3). We note that the sharp sawtooth pattern is strongest on small scales, and systematic CTE deterioration thus limits the resolution of mass reconstruction from HST-based observations of weak lensing, at a level only just below the statistical limit set by the finite number density of resolved galaxies.

Photometric redshift measurement. Extensive follow-up observations of the COSMOS field with the Subaru, Canada–France–Hawaii (CFH), Cerro Tololo Inter-American Observatory (CTIO) 4 m and the Kitt Peak National Observatory (KPNO) 4 m telescopes[27] have provided optical and near-infrared imaging in 15 bands ranging from u^* to K_s. Such multicolour data constitutes low-resolution spectroscopy, and we use a bayesian template-fitting method[28,29] to estimate the redshift and stellar mass of each galaxy (Supplementary Fig. 5). The depth of the follow-up observations ensures completeness in stellar mass down to $7 \times 10^9\ M_{\rm Sun}$ at $z < 1.05$ (ref. 11). In the source galaxy redshift distribution (Fig. 2), the foreground peaks below $z = 1.2$ all correspond to known structures in the COSMOS field.

For galaxies fainter than those with an AB magnitude of 24.5 in the F814W filter, a characteristic degeneracy exists between galaxies at $0.1 < z < 0.3$ and $1.5 < z < 3.2$ without real spectroscopy, owing to confusion between the 4,000 Å break and coronal absorption features[27]. For the purposes of weak lensing, this degeneracy is not symmetric. Distant galaxies are viewed after significant distortion and, if placed erroneously at low redshift, would create spurious power in the nearby Universe that echoes more distant structures. On the other hand, very nearby galaxies are almost unlensed so, if placed incorrectly at high redshift, they merely dilute the signal. To deal with this redshift degeneracy, we study the joint redshift probability distribution function for each galaxy. For those with best-fitting redshifts $z < 0.4$, if any probability exists above $z > 1.5$, we move the galaxy to the weighted mean of the redshift probability integrated above $z = 1.5$. This places all uncertain galaxies in the same place, and contracts two problems into the less troublesome one. We then statistically estimate the overpopulation of high-redshift slices by comparing their apparent density of galaxies to that expected from the known galaxy luminosity function[30] (Supplementary Fig. 6).

Received 14 November; accepted 30 November 2006.
Published online 7 January 2007.

1. Zwicky, F. Die Rotverschiebung von extragalaktischen Nebeln. *Helv. Phys. Acta* **6**, 110–127 (1933).
2. Bergstrom, L. Nonbaryonic dark matter: observational evidence and detection methods. *Rep. Prog. Phys.* **63**, 793–841 (2003).
3. Clowe, D. *et al.* A direct empirical proof of the existence of dark matter. *Astrophys. J.* **648**, L109–L113 (2006).
4. Blandford, R., Saust, A., Brainerd, T. & Villumsen, J. The distortion of distant galaxies by large-scale structure. *Mon. Not. R. Astron. Soc.* **251**, 600–627 (1991).
5. Kaiser, N. Weak gravitational lensing of distant galaxies. *Mon. Not. R. Astron. Soc.* **388**, 272–286 (1992).
6. Bartelmann, M. & Schneider, P. Weak gravitational lensing. *Phys. Rep.* **340**, 291–472 (2001).
7. Refregier, A. Weak gravitational lensing by large-scale structure. *Annu. Rev. Astron. Astrophys.* **41**, 645–668 (2004).
8. Davis, M., Efstathiou, G., Frenk, C. & White, S. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophys. J.* **292**, 371–394 (1985).
9. Springel, V. *et al.* Simulating the joint evolution of quasars, galaxies and their large-scale distribution. *Nature* **435**, 629–636 (2005).
10. Dekel, A. & Lahav, O. Stochastic nonlinear galaxy biasing. *Astrophys. J.* **520**, 24–34 (1999).
11. Scoville, N. *et al.* COSMOS: Hubble Space Telescope observations. *Astrophys. J.* (in the press).
12. Leauthaud, A. *et al.* COSMOS: ACS galaxy catalog. *Astrophys. J.* (submitted).
13. Schneider, P., van Waerbeke, L. & Mellier, Y. B-modes in cosmic shear from source redshift clustering. *Astron. Astrophys.* **389**, 729–741 (2002).

©2007 Nature Publishing Group

14. Hasinger, G, *et al.* The XMM-Newton wide-field survey in the COSMOS field: I. Survey description. *Astrophys. J.* (in the press).
15. Guzzo, G. *et al.* A large-scale stucture at $z=0.73$ and the relation of galaxy morphologies to local environment. *Astrophys. J.* (submitted).
16. Shen, J., Abel, T., Mo, H. & Sheth, R. An excursion set model of the cosmic web: the abundance of sheets, filaments, and halos. *Astrophys. J.* **645,** 783–791 (2006).
17. Massey, R. *et al.* Weak lensing from space: II. Dark matter mapping. *Astron. J.* **127,** 3089–3101 (2004).
18. Bacon, D. & Taylor, A. Mapping the 3D dark matter potential with weak shear. *Mon. Not. R. Astron. Soc.* **344,** 1307–1326 (2003).
19. Taylor, A. *et al.* Mapping the 3D dark matter with weak lensing in COMBO-17. *Mon. Not. R. Astron. Soc.* **353,** 1176–1196 (2003).
20. Saunders, W. *et al.* Density and velocity fields from the PSCz survey. In *Towards an Understanding of Cosmic Flows* (eds Courteau S., Strauss M. & Willick, J) *ASP Conf. Ser.,* **201,** 228–236 (Astronomical Society of the Pacific, San Francisco, 1999).
21. Rhodes, J., Refregier, A. & Groth, E. Weak lensing measurements: a revisited method and application to Hubble Space Telescope images. *Astrophys. J.* **536,** 79–100 (2000).
22. Kaiser, N. & Squires, G. Mapping the dark matter with weak gravitational lensing. *Astrophys. J.* **404,** 441–450 (1993).
23. Starck, J.-L. & Murtagh, F. *Astronomical Image and Data Analysis* 2nd edition (Springer, Heidelberg, 2006).
24. Starck, J.-L., Pires, S. & Refregier, A. Weak lensing mass reconstruction using wavelets. *Astron. Astrophys.* **451,** 1139–1150 (2006).
25. Hopkinson, G., Dale, C. & Marshall, P. Proton effects in charge-coupled devices. *IEEE Trans. Nucl. Sci.* **43,** 614–627 (1996).
26. Rhodes, J. *et al.* Removing the effects of the Advanced Camera for Surveys point spread function. *Astrophys. J.* (submitted).
27. Capak, P. *et al.* Photometric redshifts of galaxies in COSMOS. *Astrophys. J.* (submitted).
28. Benítez, N. Bayesian photometric redshift estimation. *Astrophys. J.* **536,** 571–583 (2000).
29. Mobasher, B. *et al.* The first release COSMOS optical and near-IR data and catalog. *Astrophys. J.* (submitted).
30. Steidel, C. *et al.* A study of star-forming galaxies in the $1.4<z<2.5$ redshift desert: overview. *Astrophys. J.* **604,** 534–550 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was based on observations with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc. (AURA). We also used data collected from: the XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA member states and NASA; the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan; the European Southern Observatory, Chile; the Kitt Peak National Observatory, the Cerro Tololo Inter-American Observatory, and the National Optical Astronomy Observatory, which are operated by AURA under cooperative agreement with the American National Science Foundation; the National Radio Astronomy Observatory, which is a facility of the American National Science Foundation operated under cooperative agreement by Associated Universities, Inc.; and the Canada-France-Hawaii Telescope operated by the National Research Council of Canada, the Centre National de la Recherche Scientifique de France and the University of Hawaii. The photometric redshifts used here were validated using spectra from the European Southern Observatory Very Large Telescope zCOSMOS survey. We gratefully acknowledge the contributions of the entire COSMOS collaboration, consisting of more than 70 scientists. We thank T. Roman, D. Taylor and D. Soderblom for help scheduling the extensive COSMOS observations; and A. Laity, A. Alexov, B. Berriman and J. Good for managing online archives and servers for the COSMOS data sets at NASA IPAC/IRSA. This work was supported by grants from NASA (to N.S. and R.M.).

Author Contributions A.K. processed the raw HST data, and J.-P.K. masked defects in the image. A.L., J.R. and R.M. catalogued the positions and shapes of galaxies. Y.T. and S.S. obtained multicolour follow-up data, which was processed and calibrated by S.S., P.C., H.McC. and H.A. P.C. determined galaxies' redshifts, and B.M. their stellar mass. N.S. constructed maps of stellar mass and galaxy density. A.F. processed the X-ray image and removed point sources. R.M. and A.R. produced the two-dimensional and tomographic mass maps; J.-L.S. and S.P. developed the wavelet filtering technique. D.B. and A.T. produced the three-dimensional mass reconstruction. J.T., A.F., R.E. and R.M. compared the various tracers of large-scale structure.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.M. (rjm@astro.caltech.edu).

©2007 Nature Publishing Group

LETTERS

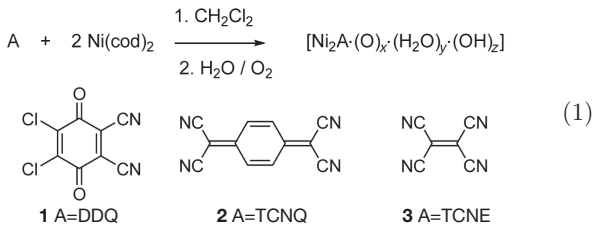
High-temperature metal–organic magnets

Rajsapan Jain[1], Khayrul Kabir[1], Joe B. Gilroy[1], Keith A. R. Mitchell[2], Kin-chung Wong[2] & Robin G. Hicks[1]

For over two decades there have been intense efforts aimed at the development of alternatives to conventional magnets, particularly materials comprised in part or wholly of molecular components[1,2]. Such alternatives offer the prospect of realizing magnets fabricated through controlled, low-temperature, solution-based chemistry, as opposed to high-temperature metallurgical routes, and also the possibility of tuning magnetic properties through synthesis. However, examples of magnetically ordered molecular materials at or near room temperature are extremely rare[3], and the properties of these materials are often capricious and difficult to reproduce. Here we present a versatile solution-based route to a new class of metal–organic materials exhibiting magnetic order well above room temperature. Reactions of the metal (M) precursor complex bis(1,5-cyclooctadiene)nickel with three different organics A—TCNE (tetracyanoethylene), TCNQ (7,7,8,8-tetracyanoquinodimethane) or DDQ (2,3-dichloro-5,6-dicyano-1,4-benzoquinone)—proceed via electron transfer from nickel to A and lead to materials containing Ni(II) ions and reduced forms of A in a 2:1 Ni:A ratio—that is, opposite to that of conventional (low Curie temperature) MA_2-type magnets. These materials also contain oxygen-based species within their architectures. Magnetic characterization of the three compounds reveals spontaneous field-dependent magnetization and hysteresis at room temperature, with ordering temperatures well above ambient. The unusual stoichiometry and striking magnetic properties highlight these three compounds as members of a class of stable magnets that are at the interface between conventional inorganic magnets and genuine molecule-based magnets.

Selected metal–cyanide-based organometallic magnets order near or above room temperature[4], although classification of these as 'molecule-based' is debatable given that the cyanide ligand is only nominally larger than the oxo (O^{2-}) bridge in conventional inorganic magnets (such as in Fe_3O_4 or CrO_2). The only molecule-based magnet which orders above room temperature[5] (transition temperature $T_c \approx 400$ K) is $V(TCNE)_x \bullet y CH_2Cl_2$ (where $x \approx 2$; $y \approx 0.5$). This pyrophoric compound is the prototypical member of a large family of low-T_c magnets based on metal (II) cations and radical anions of TCNE in a 1:2 ratio[6]. The proposed structural model is a three-dimensional coordination network[6,7] in which magnetic cooperativity arises from exchange interactions between localized spins on the metal ion and spins on the TCNE radical anion, although structural proof is still lacking. Other vanadium-based materials VR_2 (where R = radical anion) have T_c values of approximately 200 K[7,8], whereas analogous $M(TCNE)_2$ materials (where M = Fe, Co, Mn, Ni) have considerably lower (<122 K) ordering temperatures[9–11]. In general metal–TCNX (where X = E, Q) materials have a long history as conductors[12,13] and as low-temperature magnets[9–11,14], nearly all of which strictly adhere to the MR_2 stoichiometric model. The only exceptions are Ni(TCNQ) and $Ni_3(TCNQ)_2$, neither of which supports magnetic ordering[15].

Here we present investigations on reactions of bis(1,5-cyclooctadiene) nickel—$Ni(cod)_2$—with the organic oxidizing agents TCNE, TCNQ and DDQ. When these reactions were carried out in the conventional 1:2 Ni:A ratio, low yields of ill-defined products were obtained. In each case, a significant amount of unreacted organic A was recovered—an observation that prompted us to explore reactions involving higher M:A ratios. Reactions involving a 2:1 Ni:A stoichiometry (the opposite of the conventional ratio) under an argon atmosphere led to immediate formation of intensely coloured and highly air-sensitive (hydrophilic/oxophilic) cloudy solutions. Subsequent controlled exposure (~2 hours) of these solutions to air led to precipitation of near-black solid materials that were isolated by filtration to afford the final products **1**–**3** in quantitative yield (see equation (1)). Attempts to characterize the nature of the 'intermediate species' ($Ni_{2x}A_x$) before air exposure—either in solution or the solid material obtained from it—were thwarted by the intermediate's high reactivity. Clearly the original 2:1 Ni:A stoichiometry precludes complete oxidation/coordination of all of the nickel centres; the air-exposure step fulfills these requirements for the remaining nickel centres during the reaction of the intermediate ($Ni_{2x}A_x$) with molecular oxygen (oxidant) and water (coordinating ligand) (see elemental analyses). Cyclooctadiene is the only material present in the filtrate and can be quantitatively recovered; elemental analyses of **1**–**3** confirm the absence of cyclooctadiene and also reveal that both the nickel and the ligands A are incorporated quantitatively into **1**–**3**.



(1)

Elemental analyses confirm the 2:1 Ni:A ratio, as well as the presence of additional oxygen-containing species—a combination of water and oxide (O^{2-}) or hydroxide (OH^-) resulting from the hydrolysis-based work-up. Therefore, the empirical compositions of **1**–**3** can best be formulated as $[Ni_2A \bullet (O)_x \bullet (H_2O)_y \bullet (OH)_z]$ with residual solvent, where *x*, *y*, *z* depend on A and are also interdependent (see Supplementary Information for specifics). Charge-balance considerations require that the total negative charge from the ligands A and OH^-, O^{2-} balance the total positive charge on the nickel ions (see below). Efforts to pinpoint these components were thwarted by the inability to distinguish between analytically equivalent oxygen-based species (for example, $[H_2O + O^{2-}]$ versus $[2OH^-]$ in **1** or $[H_2O + O^{2-} + OH^-]$ versus $[3OH^-]$ in **2** and **3**). Powder X-ray diffraction experiments reveal all samples to be amorphous.

[1]Department of Chemistry, University of Victoria, PO Box 3065 STN CSC, Victoria, British Columbia V8W 3V6, Canada. [2]Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada.

©2007 Nature Publishing Group

The scanning transmission electron microscopy (STEM) images show all three samples to consist of irregularly shaped and sized 'grains' in the micrometre-to-submicrometre regime, with a texture like broken mud (Supplementary Information). Energy-dispersive X-ray microanalyses corroborated the elemental compositional makeup. Characteristic Ni-Kα, Lαβ X-rays were observed at 7.477 and 0.849 keV respectively[15]. The C, O and Cl Kα-peaks featured at 0.282, 0.523 (ref. 16) and 2.622 keV respectively. The presence of the oxygen peak in **2** and **3** supports the idea of the incorporation of oxygen-containing species from the hydrolysis-based work-up. Studies of several individual grains confirm consistent and uniform composition on this length scale. TEM images show smaller textural features estimated to be 30 ± 3 nm, but none of the images are reminiscent of well-defined nanostructures.

The infrared spectra of the Ni_2TCNQ- and Ni_2TCNE-based materials (**2** and **3,** respectively) are consistent with these organics present as radical anions: broad ν_{CN} bands were centred at 2,191 cm^{-1} in **2** and at 2,205 and 2,179 cm^{-1} in **3** (ref. 17). For **2** this is further corroborated by a $\delta_{C\text{-}H}$ feature at 824 cm^{-1} and $\nu_{C=C}$ stretches at 1,505 cm^{-1}, both traits of the π-delocalized $TCNQ^{-\bullet}$ moiety (ref. 14 and references therein). In contrast, the infrared spectrum of Ni_2DDQ (**1**) possesses a large bathochromic shift (> 240 cm^{-1}) in its ν_{CO} compared to neutral DDQ, which is consistent with dianionic DDQ[18]. The spectra of **1–3** are consistent with the ligands A being coordinated to the nickel ions because the ν_{CN} or ν_{CO} values do not correspond to non-coordinated TCNE and TCNQ (ref. 19 and references therein) or DDQ[19,20] in any of their normal oxidation states, as is observed in their electron-transfer salts. The infrared characteristics differ from $Ni(TCNX)_2$ materials[9,11,14,21] and all exhibit 'clean', singular CN stretches, precluding the presence of degraded forms of the ligands, which would possess their own characteristic stretches. The solid-state electronic spectra corroborate DDQ as a dianion in **1** (ref. 20) through a characteristic absorption at ~400 nm, whereas **2** and **3** have low energy absorptions consistent with the presence of the $TCNQ^{-\bullet}$ and $TCNE^{-\bullet}$ moieties respectively[22].

The oxidation state of nickel was probed by X-ray photoelectron spectroscopy. The binding energies found for the Ni $2p_{3/2}$ peaks in **1–3** are 856.5, 856.4 and 856.3 eV respectively, values in the expected range for chemically stable Ni(II) centres bound to organo- moieties[15]; ruling out nickel (0), for which the Ni $2p_{3/2}$ peaks are typically around 853 eV (ref. 23). Charge-balance considerations therefore require anionic components additional to the organic anion A^{n-}, probably metal-coordinated OH^- or O^{2-}, supported by elemental analyses.

Compounds **1–3** are readily attracted in bulk to external magnets, under ambient conditions. Quantitative magnetic studies were performed with a superconducting quantum interference device (SQUID) magnetometer. Figure 1 presents the field-cooled temperature dependence of magnetization, the $M(T)$ profiles, for **1–3**. Critical temperatures (T_c values) were estimated to be 405 K (**1**), 480 K (**2**) and 440 K (**3**) from the onset points ($M(T)\rightarrow 0$), derived by extrapolation[6–8,11]. These T_c values are more than one order of magnitude larger than those reported for $Ni(TCNX)_2$-based magnets[9,11,14,21] and are clearly dependent on the organics A.

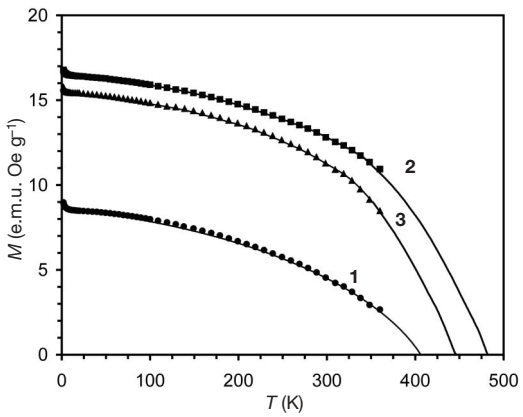


Figure 1 | Temperature dependence of field-cooled (25 Oe) magnetization for 1–3. Solid lines are extrapolations.

Table 1 | Magnetic parameters for 1–3

Compound	T_c (K)	H_c (Oe)	M_r (e.m.u. Oe g^{-1})	M_{sat} (e.m.u. Oe g^{-1})
1	405	12.9	0.46	32.2
2	480	14.0	5.49	68.6
3	440	12.4	3.69	52.9

Magnetization versus field plots at 300 K reveal that all three compounds exhibit spontaneous magnetization and distinctive hysteresis (Fig. 2a, b), substantiating long-range bulk magnetic ordering at ambient temperature. Spontaneous magnetization behaviour (Fig. 2b), expressed in gram-magnetization (employed because of unknown structure), in the field-dependent plots was typical of real magnets. Qualitatively, all three samples have visibly different attractive responses to an external magnet. Coercive fields (H_c) and remanent magnetization (M_r) of **1–3** at 300 K are tabulated (Table 1). Interestingly, the remanence of **1**, which contains a spin-inactive DDQ dianion, is an order of magnitude smaller than that of **2** or **3**, both of which contain spin-active ($S=1/2$) radical anions. On the basis of the observed coercive fields (10–15 Oe), all three compounds are soft magnets[3]. The temperature dependence of the M_{sat} values for isothermals between 1.8–100 K, and particularly their shapes in the M versus H plots, (Supplementary Information) can be attributed to a glassy magnetic state[14,24] (see below) owing to spin disorder or frustration inherent in these materials, which is less prominent at higher temperatures (>100 K) approaching or above freezing temperatures. Unexpectedly, however, the M_{sat} values abnormally increase in the low-temperature $M(H)$ measurements. The possibility that the overall magnetization may consist of contributions from two different magnetic regimes—one magnetically ordered and one consisting of non-correlated spins—cannot be ruled out, although the disordered

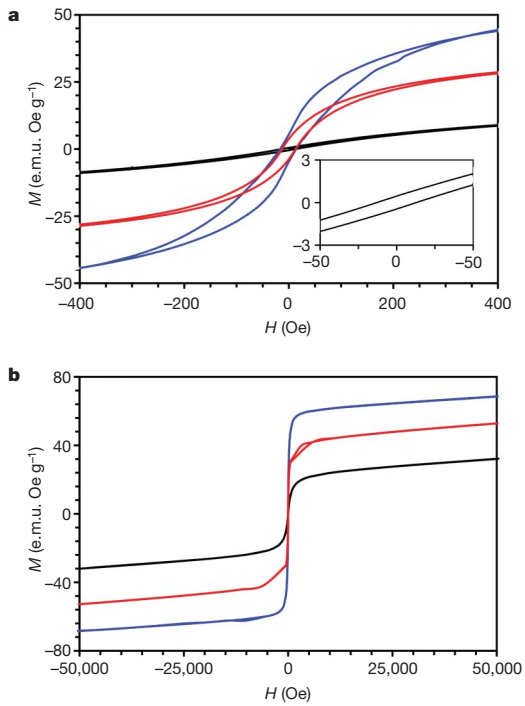


Figure 2 | Magnetic hysteresis loops for 1 (black line), 2 (blue line), and 3 (red line) at 300 K. a, ±400 Oe. Inset shows expansion of the loop for **1. b**, ±50,000 Oe.

©2007 Nature Publishing Group

non-correlated spin-glass $V(TCNE)_2 \bullet 0.5CH_2Cl_2$ molecule-based magnet[25] shows similar trends in M_{sat}. This behaviour reflects a deviation from properties reminiscent of conventional magnets.

Alternating current (a.c.) susceptibility measurements also support magnetic ordering well above room temperature. Figure 3 presents a.c. data for **3** as a representative example. Both $\chi'(T)$ (in-phase) and $\chi''(T)$ (out-of-phase) plots are frequency-dependent with peak (freezing) temperatures shifting to higher values with increasing frequencies; this trend in the χ' data is suggestive of glassy magnetic behaviour[14]. For a given frequency, the lack of multiple peaks supports the existence of a single magnetic phase[21]. The incomplete saturation in the $M(H)$ plots is also consistent with the postulated glassy state[24], although degrees of glassiness can vary as a function of crystallographic disorder[24], spin frustration arising from specific spin topology[14], residual solvent and temperature[25].

Electron transfer chemistry between $Ni(cod)_2$ and A appears to be a key requirement in the formation of **1**–**3**. Reactions of $Ni(cod)_2$ with poor oxidizing agents like 1,4-benzoquinone and 1,4-naphthoquinone lead to low yields of products that do not exhibit magnetic ordering. The corresponding reaction with duroquinone yields non-magnetic discrete π-sandwich complexes[26] and no reaction occurs with p-anthraquinone. The use of coordinating solvents like tetrahydrofuran or acetonitrile thwarts the formation of **1**–**3**, perhaps through solvent coordination to nickel[25]. Conceivably, the dissociation of cyclooctadiene with pre-coordination between the organic acceptor and Ni leads to an intermediate which facilitates electron transfer[7,8].

The X-ray photoelectron spectroscopy studies (see above) rule out sources of nickel (0) (whether bulk nickel ($T_c = 627$ K) or nickel nanoparticles, which are superparamagnetic, with blocking temperatures dependent upon surface species and size/aggregation effects[27]) as the dominant source of the magnetism. The use of $Ni(cod)_2$ to produce Ni nanoparticles requires special conditions (for example, sonochemical decomposition[28]) that afford different types of magnetic materials[29]. Pure nickel oxide (NiO) and nickel hydroxide ($Ni(OH)_2$) are antiferromagnets ($T_N = 523$ K and 24.8 K respectively) and therefore cannot justify the origin of the magnetism in **1**–**3.** Nickel oxide nanoparticles are dissimilar to **1**–**3** because their properties are dependent upon finite size effects and fit behaviour defined as superantiferromagnetic or superparamagnetic[30]. In contrast, the magnetic properties of **1**–**3** are reproducible from sample to sample without any discernible size-dependence.

A structural model for these materials is still lacking, but the compositional and magnetic data argue against the model developed for the MA_2 magnets[6–8], and the existence of several NiA_2 materials with T_c values < 50 K suggest that the materials described here have a different (and new) structural basis. The spectroscopic features suggest that the organic ligands are σ-bound (via the nitrile or carbonyl groups) to the nickel. The distinct magnetic properties of **1**, **2** and **3** strongly suggest that the organic anions are not merely present to achieve charge compensation but are integral to the magnetic structure of these materials.

The predictive value of the redox and coordination chemistry suggests that it should be possible to produce a wider range of analogous magnets by careful consideration of other organic ligands, and that the intriguing high-temperature magnetic properties of **1**–**3** could be further modified.

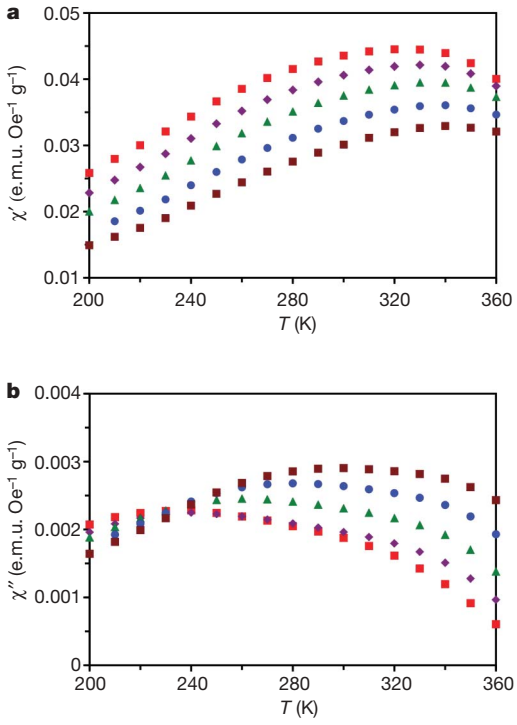


Figure 3 | Temperature dependence of the a.c. susceptibility of 3 at different frequencies. a, In-phase (χ'); **b**, out-of-phase (χ''). **0.1 Hz, red squares; 1.0 Hz, purple diamonds; 10.0 Hz, green triangles; 117.04 Hz, blue circles; 946.97 Hz, brown squares.**

METHODS

See the Supplementary Information for general information on spectroscopy (Fourier transform infrared, ultraviolet–visible, XPS), magnetic and scanning TEM equipment/measurements, and for spectroscopic, EDX, XPS, scanning TEM images, elemental analyses, and magnetic characterization data for **1**–**3**.

General. Bis(1,5-cyclooctadiene) nickel (0) {$Ni(cod)_2$} was purchased from Strem Chemicals, stored in a glove-box freezer (−30 °C) and used as received. All glassware was thoroughly pre-dried in an oven at 170 °C before use. Dichloromethane was freshly distilled under argon before use. DDQ, TCNQ and TCNE were purchased from Aldrich and also used as received.

Synthesis. In a representative experiment performed on a Schlenk line, we added 0.337 g (1.225 mmol) $Ni(cod)_2$ as a solid (weighed in dry box and taken out in a tightly sealed vial just before adding) to a vigorously stirred solution of 0.125 g (0.612 mmol) of TCNQ dissolved in 100 ml dichloromethane (DCM), under a high flow of argon at room temperature. An immediate and intense colour change from yellow to deep-green/black was observed upon mixing and after a few minutes the solution became cloudy/opaque to transmitted light. The reaction flask was immediately secured with a septum and the reaction was stirred for an additional 3.5 h while reducing the flow of argon to a normal rate. Following this, the reaction mixture was exposed to air for ~2 h and stirred at a moderate rate. The solution was then suction filtered using a normal water-jet aspirator and the precipitate was thoroughly washed with DCM, and dried in vacuum to produce a dark-coloured solid (**2**) with a quantitative yield of 0.237 g. For **1**; A = DDQ (0.125 g, 0.551 mmol); $Ni(cod)_2$ (0.303 g, 1.102 mmol). Yield of **1**: 0.230 g. For **3**; A = TCNE (0.100 g, 0.781 mmol); $Ni(cod)_2$ (0.430 g, 1.563 mmol). Yield of **3**: 0.242 g

Received 25 January; accepted 13 November 2006.

1. Kahn, O. *Molecular Magnetism* (VCH, New York, 1993).
2. Miller, J. S. Organometallic- and organic-based magnets: New chemistry and new materials for the new millennium. *Inorg. Chem.* **39,** 4392–4408 (2000).
3. Miller, J. S. & Epstein, A. J. Molecule-based magnets—an overview. *MRS Bull.* **25,** 21–28 (2000).
4. Ferlay, S., Mallah, T., Ouahes, R., Veillet, P. & Verdaguer, M. A room-temperature organometallic magnet based on Prussian blue. *Nature* **378,** 701–703 (1995).
5. Manriquez, J. M., Yee, G. T., McLean, R. S., Epstein, A. J. & Miller, J. S. A room temperature molecular/organic based magnet. *Science* **252,** 1415–1417 (1991).
6. Miller, J. S. & Epstein, A. J. Tetracyanoethylene-based organic magnets. *Chem. Commun.* 1319–1325 (1998).
7. Fitzgerald, J. P., Kaul, B. B. & Yee, G. T. Vanadium [dicyanoperfluorostilbene]$_2$·yTHF: a molecule-based magnet with $T_C \approx 205$ K. *Chem.* Commun.49–50 (2000).
8. Vickers, E. B., Selby, T. D. & Miller, J. S. Magnetically ordered (T_C=200 K) bis(tetracyanopyrazine)vanadium, $V[TCNP]_2 \cdot yCH_2Cl_2$. *J. Am. Chem. Soc.* **126,** 3716–3717 (2004).
9. Zhang, J. *et al.* [M-II(tcne)$_2$].xCH_2Cl_2 (M=Mn, Fe, Co, Ni): molecule-based magnets with T_C values above 100K and coercive fields up to 6500Oe. *Angew. Chem. Int. Edn Engl.* **37,** 657–660 (1998).

©2007 Nature Publishing Group

10. Pokhodnya, K. I., Petersen, N. & Miller, J. S. Iron pentacarbonyl as a precursor for molecule-based magnets: Formation of $Fe[TCNE]_2$ (T_C=100 K) and $Fe[TCNQ]_2$ (T_C=35 K) magnets. *Inorg. Chem.* **41,** 1996–1997 (2002).
11. Vickers, E. B., Senesi, A. & Miller, J. S. $Ni[TCNE]_2{\cdot}zCH_2Cl_2$ (T_C=13 K) and $V_xNi_{1-x}[TCNE]_y{\cdot}zCH_2Cl_2$ solid solution room temperature magnets. *Inorg. Chim. Acta* **357,** 3889–3894 (2004).
12. Kathirgamanathan, P. & Rosseinsku, D. R. Electrocrystallized metal-tetracyanoquinodimethane salts with high electrical conductivity. *J. Chem. Soc. Chem.* Commun.840–841 (1980).
13. Ferraro, J. R. & Williams, J. M. *Introduction to Synthetic Electrical Conductors* (Academic Press, New York, 1987).
14. Clerac, R. *et al.* Glassy magnets composed of metals coordinated to 7,7,8,8-tetracyanoquinodimethane: $M(TCNQ)_2$ (M = Mn, Fe, Co, Ni). *Chem. Mater.* **15,** 1840–1850 (2003).
15. Long, G. & Willett, R. D. Electrochemical synthesis, characterization and magnetic studies of Ni/TCNQ salts. *Inorg. Chim. Acta* **313,** 1–14 (2001).
16. Osan, J. *et al.* Evaluation of energy-dispersive x-ray spectra of low-Z elements from electron-probe microanalysis of individual particles. *X-ray Spectrom.* **30,** 419–426 (2001).
17. Kaim, W. & Moscherosch, M. The coordination chemistry of TCNE, TCNQ, and related polynitrile pi-acceptors. *Coord. Chem. Rev.* **129,** 157–193 (1994).
18. Miller, J. S. & Dixon, D. A. Dianion stabilization by $M(C_5(CH_3)_5)_2)^+$—theoretical evidence for a localized ring in $(DDQ)^{2-}$. *Science* **235,** 871–873 (1987).
19. Murphy, V. J. & O'Hare, D. Synthesis and magnetic characterization of $Fe(\eta^5\text{-}C_9Me_7)_2{\cdot}^+A^{\bullet-}$ (A = TCNE, TCNQ, DDQ)—X-ray structure of $Fe(\eta^5\text{-}C_9Me_7)_2{\cdot}^+TCNQ^{\bullet-}$. *Inorg. Chem.* **33,** 1833–1841 (1994).
20. Miller, J. S. *et al.* Radical ion salts of 2,3-dichloro-5,6-dicyanobenzoquinone and metallocenes—a reexamination of their magnetic and spectroscopic properties. *J. Am. Chem. Soc.* **108,** 4459–4466 (1986).
21. Vickers, E. B., Giles, I. D. & Miller, J. S. M[TCNQ]x-based magnets (M = Mn, Fe, Co, Ni; TCNQ=7,7,8,8-tetracyano-p-quinodimethane). *Chem. Mater.* **17,** 1667–1672 (2005).
22. Bell, S. E. *et al.* Novel and stable metal-metal-bonded diruthenium(I) complexes containing $TCNX^{0/\bullet-}$ in both the inner and the outer coordination sphere (TCNX = TCNE, TCNQ)—a combined EPR/ENDOR spectroscopic, UV/vis/near-IR spectroscopic, and IR spectroscopic and electrochemical investigation. *Inorg. Chem.* **31,** 3269–3276 (1992).
23. Yang, D. Q. & Sacher, E. Interaction of evaporated nickel nanoparticles with highly oriented pyrolytic graphite: Back-bonding to surface defects, as studied by x-ray photoelectron spectroscopy. *J. Phys. Chem. B* **109,** 19329–19334 (2005).
24. Buschmann, W. E., Ensling, J., Gütlich, P. & Miller, J. S. Electron transfer, linkage isomerization, bulk magnetic order, and spin-glass behavior in the iron hexacyanomanganate Prussian blue analogue. *Chem. Eur. J.* **5,** 3019–3028 (1999).
25. Pokhodnya, V. I., Pejakovic, D., Epstein, A. J. & Miller, J. S. Effect of solvent on the magnetic properties of the high-temperature $V[TCNE]_x$ molecule-based magnet. *Phys. Rev. B* **63,** 174408 (2001).
26. Schrauzer, G. N. & Thyret, H. Novel "sandwich" compounds of nickel(0); duroquinone-nickel(0) complexes with cyclic dienes. *Z. Naturforsch.* **17b,** 73–76 (1962).
27. Leslie-Pelecky, D. L. & Rieke, R. D. Magnetic properties of nanostructured materials. *Chem. Mater.* **8,** 1770–1783 (1996).
28. Koltypin, Y. *et al.* Encapsulation of nickel nanoparticles in carbon obtained by the sonochemical decomposition of $Ni(C_8H_{12})_2$. *Chem. Mater.* **11,** 1331–1335 (1999).
29. Cordente, N. *et al.* Synthesis and magnetic properties of nickel nanorods. *Nano Lett.* **10,** 565–568 (2001).
30. Khadar, M. A., Biju, V. & Inoue, A. Effect of finite size on the magnetization behavior of nanostructured nickel oxide. *Mater. Res. Bull.* **38,** 1341–1349 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Natural Sciences and Engineering Research Council of Canada and the University of Victoria for support of this research. We thank F. Palacio (Zaragoza) for assistance with some magnetic measurements and B. Gowen (Victoria) for assistance in scanning TEM and EDX measurements.

Author Contributions R.J. and R.G.H contributed equally to this work. R.J. initiated the work under the supervision of R.G.H. R.J. performed all the syntheses, control experiments, characterization, SEM and EDX measurements of the magnets. K.K. attempted to crystallize these materials, assisted in some control experiments and also processed the magnetic, X-ray photoelectron spectroscopic data. J.B.G. performed some magnetic measurements. K.A.R.M. and K.W. conducted the X-ray photoelectron spectroscopic studies. R.J. and R.G.H. wrote the manuscript. All authors have discussed the results presented and commented on the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.G.H. (rhicks@uvic.ca).

©2007 Nature Publishing Group

LETTERS

Reversible stress softening of actin networks

Ovijit Chaudhuri[1]*, Sapun H. Parekh[1]* & Daniel A. Fletcher[1]

The mechanical properties of cells play an essential role in numerous physiological processes. Organized networks of semiflexible actin filaments determine cell stiffness and transmit force during mechanotransduction, cytokinesis, cell motility and other cellular shape changes[1–3]. Although numerous actin-binding proteins have been identified that organize networks, the mechanical properties of actin networks with physiological architectures and concentrations have been difficult to measure quantitatively. Studies of mechanical properties *in vitro* have found that crosslinked networks of actin filaments formed in solution exhibit stress stiffening arising from the entropic elasticity of individual filaments or crosslinkers resisting extension[4–8]. Here we report reversible stress-softening behaviour in actin networks reconstituted *in vitro* that suggests a critical role for filaments resisting compression. Using a modified atomic force microscope to probe dendritic actin networks (like those formed in the lamellipodia of motile cells), we observe stress stiffening followed by a regime of reversible stress softening at higher loads. This softening behaviour can be explained by elastic buckling of individual filaments under compression that avoids catastrophic fracture of the network. The observation of both stress stiffening and softening suggests a complex interplay between entropic and enthalpic elasticity in determining the mechanical properties of actin networks.

Monomers of actin assemble into polar filaments that are organized by various actin-binding proteins into branched, bundled and/or crosslinked networks essential for basic cellular functions[1]. In crawling cells, growth of actin filament networks characterized by a dendritic architecture—highly branched structures with short filaments (~0.1–1 μm) oriented in the direction of migration—generates force at the cell periphery for membrane protrusions[1,9,10].

Actin filaments, as well as other biological and synthetic polymers, are categorized by the relationship between their persistence length L_p and contour length L_c. The persistence length is defined as the average length over which the filament orientation changes due to thermal fluctuations, and the contour length is the length of the completely extended filament. For flexible polymers ($L_c \gg L_p$) the resistance to extension and compression is determined by the conformational entropy of the chain and is described as entropic elasticity. Flexible polymers exhibit stress stiffening near full extension because there is ultimately only one fully extended conformation, assuming inextensibility[11]. For stiff polymers ($L_c \ll L_p$) resistance to extension, bending and compression is due to straining of molecular links from equilibrium, which is quantified by the bending modulus κ and described as enthalpic elasticity. Under compressional forces, stiff polymers buckle at the Euler buckling force[12] $F_b = \pi^2\kappa/L_c^2$, whereas there is no equivalent buckling instability for flexible polymers because random thermal forces exceed the Euler buckling force. Actin filaments are considered to be semiflexible polymers because their persistence length (~10–17 μm)[8,13] is comparable to their physiological contour length (~0.1–10 μm)[10]. At physiological temperatures, individual filaments are expected to exhibit a combination of both entropic and enthalpic elasticity that is sensitive to the ratio of L_c to L_p and stiffen at strains much lower than in flexible polymers[6,14].

The mechanical properties of actin networks have been studied extensively using rheology of random networks formed by mixing purified actin and various actin-binding proteins in solution[2,4–6]. Recent studies have demonstrated stress stiffening of crosslinked actin networks *in vitro* under shear and explained this behaviour as arising from the inherent stress stiffening of individual filaments or flexible crosslinkers between filaments under extension, suggesting that the elasticity of these networks is entropic[5–7]. Filaments resisting compression have not been found to be important[6]. However, as filament length decreases from lengths seen in these *in vitro* assays (~2–70 μm) to physiological values, filaments are expected to support more significant bending and compressional forces[14]. Indeed, the increasing importance of enthalpic elasticity has been predicted for high concentrations of actin and crosslinkers when the distance between crosslinks becomes small[5,15]. Here we report that dendritic actin networks reconstituted *in vitro* exhibit a regime of reversible stress softening. This stress softening can be explained by the elastic buckling of individual filaments, providing evidence for an elastic response of dendritic actin networks that is enthalpic under large compressional forces and is dominated by the resistance of individual filaments to compression.

We studied the mechanical properties of growing dendritic actin networks *in vitro* using a recently developed dual-cantilever atomic force microscope (AFM) that uses a second cantilever to improve instrument stability (see Supplementary Information A, B)[16]. Dendritic networks, like those formed during cell motility, must be assembled hierarchically from a nucleating surface and cannot be formed randomly in equilibrium solutions of component proteins. In our assay, the nucleation promotion factor ActA, from the bacterial pathogen *Listeria monocytogenes,* was nonspecifically adhered to the end of an AFM cantilever before immersion in cytoplasmic extract from *Xenopus laevis* eggs. Upon immersion, ActA activates the Arp2/3 complex, which nucleates growing filaments as 70° branches off of existing filaments, catalysing the formation of a growing dendritic actin network between the cantilever and a nearby glass surface (Fig. 1a,b)[1,10,17]. Surfaces coated with ActA and other nucleation promotion factors have been shown to grow dendritic actin networks in cytoplasmic extract[17,18]. The AFM cantilever behaves like a hookean spring for small deflections (for which $F = k\Delta x$), allowing measurement and application of compressional forces to the growing network.

Using AFM-based microrheology[19,20] (Fig. 1c, d, Methods), we measured frequency-dependent elastic (storage) and viscous (loss) moduli. After normalization at a reference frequency (see Supplementary Information C), we find that the frequency dependence of the elastic modulus E' is consistent with power-law rheology, $E' \propto f^x$, with $x = 0.13$ (Fig. 2). This falls within the range of exponents measured from cells ($x = 0.12$–0.25)[20–22]. The average linear elasticity of the dendritic actin networks, 985 ± 655 Pa (mean ± s.d.)

[1]UC San Francisco /UC Berkeley Joint Graduate Group in Bioengineering and Department of Bioengineering, University of California at Berkeley, Berkeley, California 94720, USA.
*These authors contributed equally to this work.

©2007 Nature Publishing Group

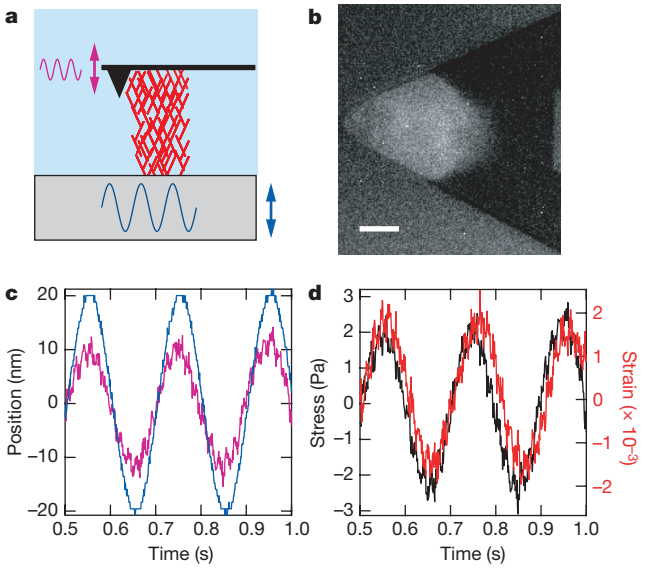


Figure 1 | AFM-based microrheology of growing dendritic actin networks. **a**, Cartoon illustrating the measurement geometry in which the surface is driven sinusoidally (blue sinusoid and double-headed arrow), and the force transmitted through the network (red mesh) is transduced by the cantilever (pink sinusoid and double-headed arrow). **b**, Fluorescence micrograph of the actin network, which is used to calculate the network area *A*. Scale bar is 10 μm. **c**, Graph showing surface drive and cantilever response signal as a function of time for a 5 Hz measurement (colours are as in **a**). Note the cantilever response is damped with respect to the drive signal indicating compression of the network. This technique has the effect of applying a sinusoidal stress on the network where hydrodynamic coupling was found to be negligible (see Supplementary Information B). **d**, Stress and strain graph calculated from measurement in **c** showing stress (black) and strain (red) as a function of time (see Methods).

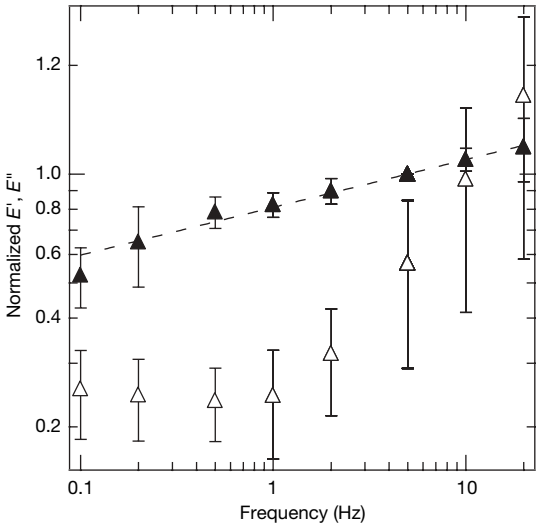


Figure 2 | Frequency dependence of elastic (filled triangles, *E′*) and viscous (open triangles, *E″*) moduli. The traces were constructed by averaging normalized data from 11 separate experiments and 21 different frequency sweeps. Each measurement of the elastic and viscous moduli was normalized by the average elastic modulus at 5 Hz taken before and after the measurement (see Supplementary Information C). The best-fit power-law exponent for $E'(f)$ was determined to be $x = 0.13$ (dotted line), and the average elastic modulus at 5 Hz was 985 ± 655 Pa (mean ± s.d.), which are consistent with previous studies on cells. In addition to the power-law behaviour, the viscous modulus has a similar shape to those seen previously. Error bars on both curves are normalized s.d.

at 5 Hz, is similar to the elastic modulus measured on various cell types[19–22] and in a previous reconstitution of actin-based motility[23]. Dendritic actin network elasticity is significantly higher than the elasticity of actin networks reconstituted in solution containing the Arp2/3 complex (~1 Pa), though differences in concentration and components could account for this disparity[24,25]. The average elasticity of the actin networks studied here was found to be independent of prestressing by myosin II motors (see Supplementary Information D).

To understand further the mechanical properties of growing dendritic actin networks, we probed the stress dependence of the elastic modulus[5,7] (see Methods). A typical experiment is shown in Fig. 3a (black trace) where stress was increased on the network incrementally, and the elasticity at each value of applied stress was measured. For stresses up to ~15 Pa the elasticity remained constant, indicating a linear elastic regime. Then the elasticity increased with stress in a stress-stiffening regime, as has been seen previously[4–6], for stresses up to a critical stress, $\sigma_c \approx 270$ Pa. Above the critical stress, we found that the elasticity of the network gradually decreased with stress in a stress-softening regime.

Stress softening has been previously explained by network rupture or crosslinker rearrangement. In rigidly crosslinked actin networks, stress softening has been attributed to the fracture of extended filaments or crosslinking/branch points at σ_c, after which elasticity drastically decreased[4–6]. Alternatively, softening was proposed to occur as a result of the unbinding of flexible crosslinkers above σ_c, which either remain unbound or re-bind to form crosslinks at different positions[7]. For either of these explanations, stress softening would reflect permanent alterations in the network that would lead to irreversibility in the elasticity of the network. That is, higher elasticities could not be recovered by reducing network loading from stresses above σ_c (refs 4–7). However, in dendritic actin networks, the stress-softening behaviour was reversible: the elasticity measured as the stress was reduced to σ_c matched the elasticity seen for increasing stress (Fig. 3a, red trace). This was seen in all experiments (Fig. 3b), so stress softening in the dendritic actin network must arise from a reversible mechanism.

A plausible explanation for reversible stress softening is through elastic buckling of individual filaments under compression. A population of filaments in the dendritic network, based on their length and orientation, will begin to buckle at a threshold stress. Upon buckling, these filaments are infinitely compliant while still supporting F_b (ref. 12). As a result, the number of load-bearing elements decreases for higher stresses, resulting in a decrease in the effective stiffness of the network. As the stress is increased, more filaments buckle, reducing the elasticity of the network further. Because filaments are assembled into an interconnected dendritic network, buckled filaments do not collapse completely, and they can unbuckle when the force is reduced, making the process of buckling reversible with load. Stress softening has been predicted from simulations of athermal crosslinked actin networks to occur as a result of filament buckling and also in an elastic element model of the cytoskeleton[26,27], although such models do not predict stress stiffening before softening. Interestingly, reversible elastic buckling of component elements is observed under high compressional forces in some types of foams[28]. Electron micrographs have shown the actin cytoskeleton ultrastructure to exhibit similarities with open lattice foams, so that this buckling behaviour might be expected[3].

Buckling of individual filaments can occur at forces consistent with the observed stress softening, based on a simple calculation. Using published electron micrographs of dendritic actin networks reconstituted *in vitro* in a similar biochemical system, we estimate filament lengths L_c to be 0.1–1 μm (ref. 17). We calculate an expected buckling force F_b of 0.5–50 pN per filament using these lengths and assuming Euler buckling, although the behaviour in a constrained environment can lead to higher buckling forces[29]. In our experiments, the average force per filament at σ_c (233 Pa, the mean value from the inset of Fig. 3b), using an average filament spacing of 50–100 nm (ref. 17), is calculated to be 0.45–2 pN, which lies within the

©2007 Nature Publishing Group

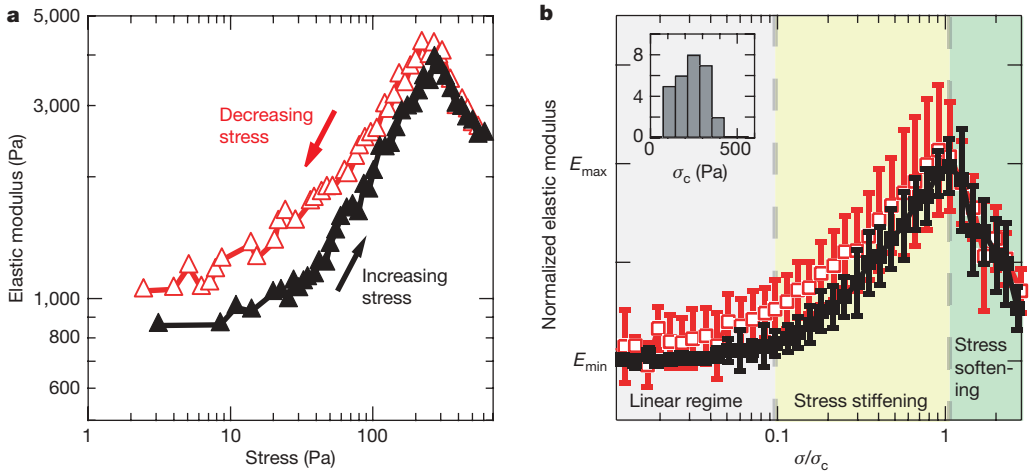


Figure 3 | Dendritic actin networks exhibit stress stiffening and reversible stress softening. **a**, In a typical nonlinear elasticity measurement, the stress on the network is first increased incrementally (black trace) to and then decreased incrementally from a maximum stress (red trace) of ~600 Pa, with the elasticity measured at each stress at 5 Hz. The elasticity remains constant for stresses up to ~15 Pa and then increases in a stress-stiffening regime. For stresses above the critical stress σ_c of ~270 Pa, the elasticity decreases in a stress-softening regime that is reversible, as indicated by the overlay of the black and red traces. **b**, Averaged and normalized trace of the nonlinear elasticity of actin networks (see Supplementary Information A). Each individual measurement was normalized by the difference between the elasticity before the measurement E_{min} and the maximum elasticity for increasing stresses E_{max} and σ_c. The results of 28 different measurements from 12 separate experiments were averaged together (mean ± s.d. shown) and found to exhibit three distinct regimes of elasticity: linear, stress stiffening and stress softening. The stress softening is shown to be reversible. Note that the elasticity in **b** is shown on a linear scale while the elasticity in **a** is shown on a log scale. The inset shows a histogram of σ_c for which the mean value was 233 Pa.

lower range of predicted buckling forces. We note that the buckling instability is smoothed entropically for a semiflexible polymer at finite temperature, so that individual polymers will undergo stress softening as the compressional force approaches the Euler buckling force. The overlap in the lower range of predicted buckling forces with the range of calculated applied force per filament at σ_c supports the idea that buckling explains stress softening, because σ_c represents the threshold stress at which filament buckling dominates nonlinear elasticity. As the stress is increased, up to $3\sigma_c$ in our experiments, shorter filaments buckle, and the elasticity decreases further.

Our measurements of nonlinear elasticity in dendritic actin networks are consistent with a model in which a combination of compression, bending and extension gives rise to network mechanical properties (Fig. 4). As stress is initially applied to the network, the elasticity increases as a result of entropic resistance to filament and flexible crosslinker extension normal to the direction of compression, in addition to possible effects from nonlinear compliance of the Arp2/3 complex (Fig. 4a, b). As stress on the network is further increased, filaments oriented in the direction of compression begin to buckle, reducing the elasticity of the network at higher stresses (Fig. 4c, d). Buckling occurs only after filaments have already been supporting a load, so the enthalpic resistance of filaments to compression is likely to play a significant role in the linear and stress-stiffening regimes.

The difference in the elasticity of dendritic networks grown from surfaces and crosslinked networks formed in solution can be explained in part by the actin concentration in the network. The modulus of elasticity is expected to scale as $E' \approx C_A^{5/2}$ (ref. 8) for isotropically crosslinked actin networks, where C_A is the concentration of actin in the network and the crosslinks are assumed to be rigid. The concentration of actin in dendritic networks has been estimated to be ~1 mM (ref. 30), whereas the concentration of networks studied *in vitro* was of the order of ~10 μM (refs 24, 25), suggesting that the magnitude of elasticities for dendritic networks should be significantly higher. Component concentration alone is sufficient to describe network properties in flexible polymer networks ($L_c \gg L_p$), where L_p is much less than the distance between crosslinks, because initial orientation and lengths of the filaments do not matter. However, for semiflexible polymer networks, filament length can be important when L_p is greater than the distance between crosslinks[15]. Additionally, the particular orientation of a filament in the network determines whether the filament deforms by compression,

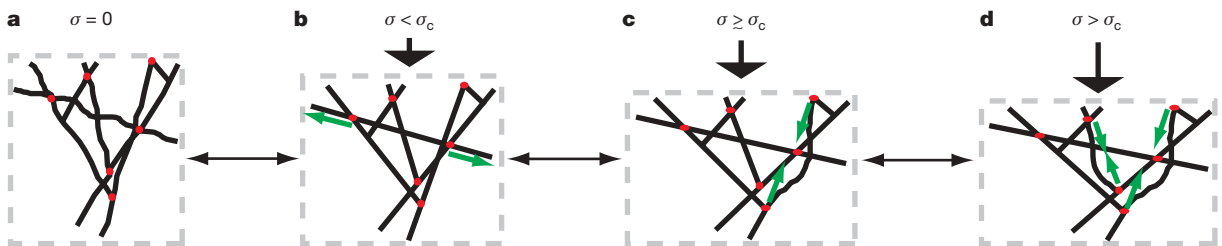


Figure 4 | Stress stiffening and stress softening can arise in dendritic networks owing to filaments resisting extension and buckling of filaments resisting compression. **a**, **b**, When the stress on the network (σ, indicated by black arrows) is increased from $\sigma = 0$, a population of filaments or crosslinkers is stretched (as indicated by green arrows) as the material expands laterally, and the resistance to extension of filaments increases owing to entropic elasticity, leading to a stress-stiffening regime. **c**, However, as the stress is increased above σ_c, some filaments resisting compression buckle when the compressional force (green arrows) exceeds the Euler buckling force. Buckled filaments exhibit infinite compliance, so they no longer contribute to the elasticity, but they do not collapse because they have connections with the network and thus still support the buckling force. **d**, As the stress is further increased, more filaments buckle and the elasticity of the network is decreased further, leading to the stress-softening regime. In principle, this process is completely reversible because buckled filaments will unbuckle once the stress is reduced.

©2007 Nature Publishing Group

bending or extension. Our finding of an elastic response in which filaments resisting extension and compression are both significant, suggests that architecture—filament length and orientation—influences the elastic behaviour of actin networks.

The reversible elastic behaviour and large elastic modulus of dendritic actin networks indicate that these networks have an architecture that is geared towards bearing high compressive loads. Because the leading edge of crawling cells is normally under compression during motility, we expect these measurements to be relevant to the mechanics of lamellipodial protrusions. Dynamic remodelling of dendritic actin network architecture may provide a mechanism for altering network elasticity in response to changing external loads.

METHODS

AFM-based microrheology. In AFM based microrheology, the deformation of the material is measured in response to a small sinusoidal stress at a given frequency, f (refs 19, 20). The elastic modulus E' and the viscous modulus E'' are then calculated as $E'(f)+iE''(f)=\tilde{\sigma}(f)/\tilde{\varepsilon}(f)$, where $\tilde{\sigma}$ and $\tilde{\varepsilon}$ are the Fourier transforms of the stress σ and the strain ε, respectively (Fig. 1d). This method is necessary for growing actin networks because only the response at the drive frequency is analysed, thereby decoupling mechanical property measurements from network growth. Changes in network length due to growth over measurement timescales (~5 min) would obscure the true strain values needed for traditional stress–strain measurements.

After the network length reached ~6 μm while growing under a small constant force, the force clamp (see Supplementary Information A) was released and the bottom surface was sinusoidally oscillated with an amplitude of 20 nm at 0.1, 0.2, 0.5, 1, 2, 5, 10 and 20 Hz using the piezoelectric positioner. This applies a sinusoidal stress to the network where hydrodynamic coupling was found to be negligible (see Supplementary Information B). At each frequency, the surface was oscillated ten times and the cantilever deflection data was used to calculate both the σ and ε (shown in Fig. 1) on the network.

Stress-dependent elasticity measurements. We adapted a technique to quantify the stress dependence of elasticity used by Gardel *et al.* on crosslinked random actin networks to our AFM-based system[5,7]. Because the relaxation processes are very slow ($x \approx 0.13$), the elasticity at a given stress changes very slowly, and this technique provides a well-defined stress–elasticity relation. In this technique, the bottom surface was incrementally stepped into (or away from) the network, and a rheology test at 5 Hz was performed to obtain the elastic modulus as a function of stress on the network. At the beginning of a nonlinear elasticity measurement, the stress on the network was reduced to nearly zero, and a 5 Hz rheology test was conducted. We incrementally increased the force on the network, and obtained the elastic modulus at each incremental stress. The stress on the network was increased until ~600 Pa, after which we incrementally decreased the stress on the network by moving the bottom surface away from the cantilever, continuing to conduct 5 Hz rheology tests after each step. Subsequent tests on the same network were started at a stress of zero.

Data analysis. In each measurement, the cantilever voltage signals were converted to deflection in Igor Pro 5 (Wavemetrics, Inc) using a voltage-deflection calibration constant that was determined at the beginning of each experiment. The length of the actin network was calculated as the distance between the two cantilevers plus an additional 3 μm due to the size of the tip. The force was determined by multiplying cantilever deflection by the stiffness (100 pN nm^{-1}). The network area was determined by thresholding the image at a greyscale value midway between the highest and lowest value along the cantilever in the ImageJ program (NIH). There is a 0.1% error in this measurement due to the diffraction-limited resolution of the fluorescence imaging.

To obtain the elastic and viscous moduli, the best-fit line was initially subtracted from the measurement cantilever deflection to eliminate the effect of slow timescale growth. We approximate the geometry of the system as a network sandwiched between two parallel plates so that $\sigma=F/A$ and $\varepsilon=\Delta L/L$, where F is the force exerted on the network by the cantilever, A is the cross-sectional area of the network and L is the length of the network. The parallel plate approximation ignores the effect of the pyramidal geometry of the tip because the tip typically represents less than 2.5% of the network area.

Additional methods are described in Supplementary Information A.

Received 28 March; accepted 16 November 2006.

1. Pollard, T. D. & Borisy, G. G. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112,** 453–465 (2003).
2. Janmey, P. A. & Weitz, D. A. Dealing with mechanics: mechanisms of force transduction in cells. *Trends Biochem. Sci.* **29,** 364–370 (2004).
3. Satcher, R. L. & Dewey, C. F. Theoretical estimates of mechanical properties of the endothelial cell cytoskeleton. *Biophys. J.* **71,** 109–118 (1996).
4. Xu, J. Y., Tseng, Y. & Wirtz, D. Strain hardening of actin filament networks—Regulation by the dynamic cross-linking protein α-actinin. *J. Biol. Chem.* **275,** 35886–35892 (2000).
5. Gardel, M. L. *et al.* Elastic behavior of cross-linked and bundled actin networks. *Science* **304,** 1301–1305 (2004).
6. Storm, C., Pastore, J. J., MacKintosh, F. C., Lubensky, T. C. & Janmey, P. A. Nonlinear elasticity in biological gels. *Nature* **435,** 191–194 (2005).
7. Gardel, M. L. *et al.* Prestressed F-actin networks cross-linked by hinged filamins replicate mechanical properties of cells. *Proc. Natl Acad. Sci. USA* **103,** 1762–1767 (2006).
8. Mackintosh, F. C., Kas, J. & Janmey, P. A. Elasticity of semiflexible biopolymer networks. *Phys. Rev. Lett.* **75,** 4425–4428 (1995).
9. Mullins, R. D., Heuser, J. A. & Pollard, T. D. The interaction of Arp2/3 complex with actin: Nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proc. Natl Acad. Sci. USA* **95,** 6181–6186 (1998).
10. Svitkina, T. M. & Borisy, G. G. Arp2/3 complex and actin depolymerizing factor cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *J. Cell Biol.* **145,** 1009–1026 (1999).
11. Bustamante, C., Marko, J. F., Siggia, E. D. & Smith, S. Entropic elasticity of λ-phage DNA. *Science* **265,** 1599–1600 (1994).
12. Landau, L. D. & Lifshitz, E. M. *Theory of Elasticity* (Butterworth-Heinemann, Oxford, 1986).
13. Gittes, F., Mickey, B., Nettleton, J. & Howard, J. Flexural rigidity of microtubules and actin-filaments measured from thermal fluctuations in shape. *J. Cell Biol.* **120,** 923–934 (1993).
14. Kroy, K. & Frey, E. Force-extension relation and plateau modulus for wormlike chains. *Phys. Rev. Lett.* **77,** 306–309 (1996).
15. Head, D. A., Levine, A. J. & MacKintosh, E. C. Deformation of cross-linked semiflexible polymer networks. *Phys. Rev. Lett.* **91,** 108102 (2003).
16. Parekh, S. H., Chaudhuri, O., Theriot, J. A. & Fletcher, D. A. Loading history determines the velocity of actin-network growth. *Nature Cell Biol.* **7,** 1119–1123 (2005).
17. Cameron, L. A., Svitkina, T. M., Vignjevic, D., Theriot, J. A. & Borisy, G. G. Dendritic organization of actin comet tails. *Curr. Biol.* **11,** 130–135 (2001).
18. Cameron, L. A., Footer, M. J., van Oudenaarden, A. & Theriot, J. A. Motility of ActA protein-coated microspheres driven by actin polymerization. *Proc. Natl Acad. Sci. USA* **96,** 4908–4913 (1999).
19. Mahaffy, R. E., Shih, C. K., MacKintosh, F. C. & Kas, J. Scanning probe-based frequency-dependent microrheology of polymer gels and biological cells. *Phys. Rev. Lett.* **85,** 880–883 (2000).
20. Alcaraz, J. *et al.* Microrheology of human lung epithelial cells measured by atomic force microscopy. *Biophys. J.* **84,** 2071–2079 (2003).
21. Stamenovic, D., Suki, B., Fabry, B., Wang, N. & Fredberg, J. J. Rheology of airway smooth muscle cells is associated with cytoskeletal contractile stress. *J. Appl. Physiol.* **96,** 1600–1605 (2004).
22. Fabry, B. *et al.* Scaling the microrheology of living cells. *Phys. Rev. Lett.* **87,** 148102 (2001).
23. Marcy, Y., Prost, J., Carlier, M. F. & Sykes, C. Forces generated during actin-based propulsion: A direct measurement by micromanipulation. *Proc. Natl Acad. Sci. USA* **101,** 5992–5997 (2004).
24. Tseng, Y. & Wirtz, D. Dendritic branching and homogenization of actin networks mediated by Arp2/3 complex. *Phys. Rev. Lett.* **93,** 258104 (2004).
25. Nakamura, F., Osborn, E., Janmey, P. A. & Stossel, T. P. Comparison of filamin A-induced cross-linking and Arp2/3 complex-mediated branching on the mechanics of actin filaments. *J. Biol. Chem.* **277,** 9148–9154 (2002).
26. Onck, P. R., Koeman, T., van Dillen, T. & van der Giessen, E. Alternative explanation of stiffening in cross-linked semiflexible networks. *Phys. Rev. Lett.* **95,** 178102 (2005).
27. Coughlin, M. F. & Stamenovic, D. A tensegrity model of the cytoskeleton in spread and round cells. *J. Biomech. Eng. Trans. Asme* **120,** 770–777 (1998).
28. Gibson, L. J. & Ashby, M. F. *Cellular Solids: Structure and Properties* (Pergamon Press, Cambridge, 1988).
29. Brangwynne, C. P. *et al.* Microtubules can bear enhanced compressive loads in living cells because of lateral reinforcement. *J. Cell Biol.* **173,** 733–741 (2006).
30. Pollard, T. D., Blanchoin, L. & Mullins, R. D. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu. Rev. Biophys. Biomol. Struct.* **29,** 545–576 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. W. Shaevitz, M. J. Rosenbluth, S. Pronk, P. L. Geissler and J. Alcaraz for discussions and reading of the manuscript as well as the entire Fletcher laboratory for support. We are also grateful to R. L. Jeng and M. J. Footer for assistance in protein preparation. This work was supported by an ASEE NDSEG Fellowship to O.C., an ARCS Fellowship to S.H.P., and an NSF Career Award and NIH grants to D.A.F.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.A.F. (fletch@berkeley.edu).

©2007 Nature Publishing Group

LETTERS

Seasonal characteristics of the Indian Ocean Dipole during the Holocene epoch

Nerilie J. Abram[1]†, Michael K. Gagan[1], Zhengyu Liu[2,3,4], Wahyoe S. Hantoro[5], Malcolm T. McCulloch[1] & Bambang W. Suwargadi[5]

The Indian Ocean Dipole[1,2] (IOD)—an oscillatory mode of coupled ocean–atmosphere variability—causes climatic extremes and socio-economic hardship throughout the tropical Indian Ocean region[1–5]. There is much debate about how the IOD interacts with the El Niño/Southern Oscillation (ENSO) and the Asian monsoon, and recent changes in the historic ENSO–monsoon relationship[6] raise the possibility that the properties of the IOD may also be evolving. Improving our understanding of IOD events and their climatic impacts thus requires the development of records defining IOD activity in different climatic settings, including prehistoric times when ENSO and the Asian monsoon behaved differently from the present day. Here we use coral geochemical records from the equatorial eastern Indian Ocean to reconstruct surface-ocean cooling and drought during individual IOD events over the past ~6,500 years. We find that IOD events during the middle Holocene were characterized by a longer duration of strong surface ocean cooling, together with droughts that peaked later than those expected by El Niño forcing alone. Climate model simulations suggest that this enhanced cooling and drying was the result of strong cross-equatorial winds driven by the strengthened Asian monsoon of the middle Holocene. These IOD–monsoon connections imply that the socioeconomic impacts of projected future changes in Asian monsoon strength may extend throughout Australasia.

The IOD, also termed the Indian Ocean Zonal Mode, involves a reversal of the sea surface temperature (SST) gradient and winds across the equatorial Indian Ocean from their climatological state. Climate simulations suggest that IOD events can be initiated by El Niño conditions that decrease the thermocline depth in the Indonesian throughflow region, and cause an eastward migration of the Pacific Walker circulation cell (which encourages atmospheric subsidence over the eastern Indian Ocean)[7,8]. Simulations also indicate that IOD events can result from processes internal to the Indian Ocean region, through intensified Hadley circulation of the Asian–Australian monsoon, which strengthens the southeasterly tradewinds and promotes upwelling of cooler water in the equatorial eastern Indian Ocean[7,8]. However, the ability to test these simulations is restricted by the spatial limitations and brevity of high-quality historical climate records (Supplementary Fig. 1) and the scarcity of proxy records[4,9], particularly in the eastern IOD region.

High-resolution reconstructions of past climate in the tropical ocean regions can be obtained through the analysis of geochemical tracers in the skeletons of massive *Porites* sp. corals. The coral reefs of the Mentawai Islands, western Indonesia, lie within the eastern Indian Ocean sector used to define the Dipole Mode Index of IOD activity[2] (Fig. 1a). During strong IOD events, upwelling in the eastern IOD sector results in anomalous cooling of SSTs and intense drought. An unambiguous record of these eastern IOD SST and rainfall anomalies can potentially be provided through coupled measurements of Sr/Ca (a proxy for SST[10]) and oxygen isotope ratios ($\delta^{18}O$; a proxy for SST and changes in rainfall/evaporation[11]) in corals from the Mentawai Islands. Furthermore, great palaeoearthquakes along the Mentawai island arc[12] have frequently uplifted the fringing reefs surrounding the Mentawai Islands, leaving individual fossil *Porites* colonies well-preserved in growth position. This

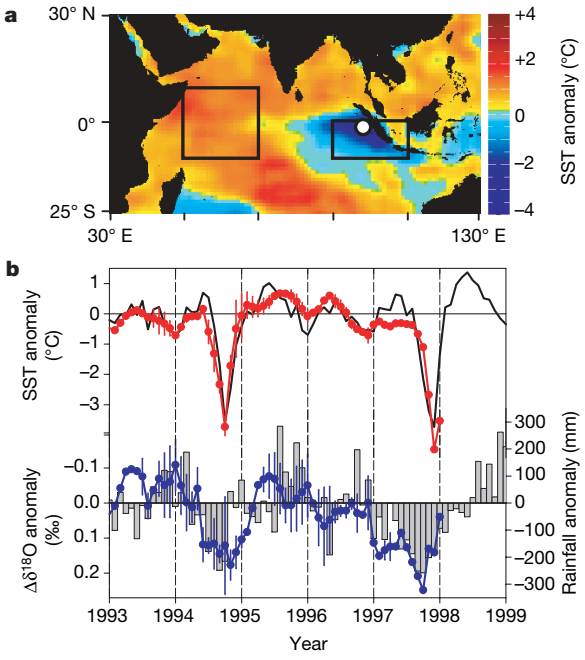


Figure 1 | IOD climate anomalies. a, SST anomalies during November of the 1997 IOD event[13], when anomalous cooling in the east and warming in the west produced a reversal of the equatorial SST gradient across the Indian Ocean. Boxes mark the eastern and western sectors used to define the Dipole Mode Index[2], and the white circle shows the location of the Mentawai Islands. **b**, In the Mentawai Islands the strong IOD events of 1994 and 1997 were characterized by cool SST anomalies (black curve)[13] and drought (grey bars)[29]. These distinct IOD SST and rainfall anomalies are preserved, respectively, in coral Sr/Ca–SST (red curve) and $\Delta\delta^{18}O$ (blue curve) anomalies[11]. Coral time series between July 1993 and February 1997 are based on the average of two coral records, with error bars showing the difference between the coral records for each monthly data point.

[1]Research School of Earth Sciences, The Australian National University, Canberra, Australian Capital Territory 0200, Australia. [2]Center for Climatic Research, University of Wisconsin-Madison, 1225 W. Dayton Street, Madison, Wisconsin 53706, USA. [3]Earth Environment Institute, Chinese Academy of Science, Xi'an 710075, China. [4]The Ocean University of China, Qingdao 266003, China. [5]Research and Development Center for Geotechnology, Indonesian Institute of Sciences (LIPI), Bandung 40135, Indonesia. †Present address: British Antarctic Survey, Natural Environment Research Council, Cambridge CB3 0ET, UK.

©2007 Nature Publishing Group

allows us to examine the ocean–atmosphere signatures of individual IOD events for slices of time since ~6,500 calibrated years before present (BP).

To test the suitability of Mentawai corals for reconstructing IOD SST and rainfall anomalies, Sr/Ca and $\delta^{18}O$ were analysed at high (~monthly) resolution in two modern corals from the Mentawai Islands (Supplementary Methods). Both corals were collected in open, fore-reef slope environments on fringing reef systems, at water depths of 1.2 m (coral TN99-A-4) and ~6 m (coral TM01-A-10; Supplementary Table 1). Both provide a record of the strong IOD event in 1994. The TM01-A-10 coral record also extends into the 1997 IOD event, when the Mentawai reefs experienced widespread coral mortality[4]. The coral Sr/Ca–SST anomalies reliably reconstruct both the timing and magnitude of IOD cooling during 1994 and 1997, as documented by blended buoy, ship and satellite SST records[13] (Fig. 1b, Supplementary Figs 1 and 2). The coral records of residual $\delta^{18}O$ ($\Delta\delta^{18}O$), determined by using the Sr/Ca–SST records to remove the temperature component of coral $\delta^{18}O$ measured for the same samples, reflect changes in the $\delta^{18}O$ composition of sea water that are due to variations in the surface–ocean balance of rainfall and evaporation[11]. Comparison with station and gridded rainfall data shows that the Mentawai coral $\Delta\delta^{18}O$ records replicate the droughts in western Indonesia that accompanied the 1994 and 1997 IOD events (Fig. 1b).

The distinct cool/dry IOD signal in the Mentawai coral Sr/Ca and $\delta^{18}O$ records indicates that corals from the eastern upwelling region can be used as an independent proxy for strong IOD activity. Further analysis of a 140-year-long coral record from the Mentawai Islands has identified two more exceptionally strong IOD signals in 1961 and 1877 AD[4] (Supplementary Fig. 3). Both years were times of extreme climate in the Indian Ocean region[2,5,14,15], consistent with strong IOD events. Indeed, 1961, 1994 and 1997 correspond with the three strongest droughts in the rainfall record of the Padang-Tabing station in western Indonesia (1° S, 100° E), which was established as part of an extensive network of Indonesian rainfall stations following the devastating drought of 1877 (ref. 15). The modern coral records suggest that strong IOD events occur in both El Niño (1877, 1994 and 1997) and non-El Niño (1961) years, consistent with climate simulation results[7,8]. IOD events also appear to exist over a wide range of monsoon conditions, with the 1877 IOD event corresponding to a year of extreme monsoon failure, while the 1961 IOD occurred during the heaviest monsoon in the nearly 200-year-long all-India monsoon rainfall index[16]. Hence, understanding how the IOD, ENSO and Asian monsoon systems interact has important implications for the predictability of climate extremes in the Indian Ocean region.

During the early to middle Holocene, broad-scale strengthening of the Asian summer monsoon occurred owing to increased Northern Hemisphere summer insolation[17]. At the same time, ENSO variability was suppressed[18,19] and a more La Niña-like mean state existed across the tropical Pacific Ocean[20]. Thus, mid-Holocene fossil corals from the Mentawai Islands provide a good opportunity to examine how externally forced increases in Asian monsoon strength (which have been predicted for the coming century[21–23]) may influence the IOD system. To do this, reconstructions of four individual IOD events from the middle Holocene (6,570–4,160 calibrated years BP; Fig. 2) were averaged to produce a composite IOD event for a period of strong Asian monsoon forcing (Fig. 3, Supplementary Fig. 3, Supplementary Table 2). For comparison, a composite IOD event was also constructed using four late Holocene IOD events (from 2,180 calibrated years BP to the present), when the Asian monsoon weakened and El Niño events were relatively strong and frequent.

Comparison of the composite SST anomalies shows that, on average, the magnitude of cooling during strong eastern IOD upwelling events is similar for the strong El Niño and strong monsoon scenarios (Fig. 3; -3.9 ± 0.2 °C and -4.3 ± 0.6 °C, respectively). The timing of peak IOD cooling also appears to have remained essentially constant with respect to the seasonal cycle of SST. Today, the timing of peak IOD cooling is constrained by the cross-equatorial wind reversal at the end of the Asian summer monsoon season (November), which brings an abrupt end to Ekman upwelling and cooling along the coast of Sumatra[1]. Thus, the similarity of the seasonal timing of IOD

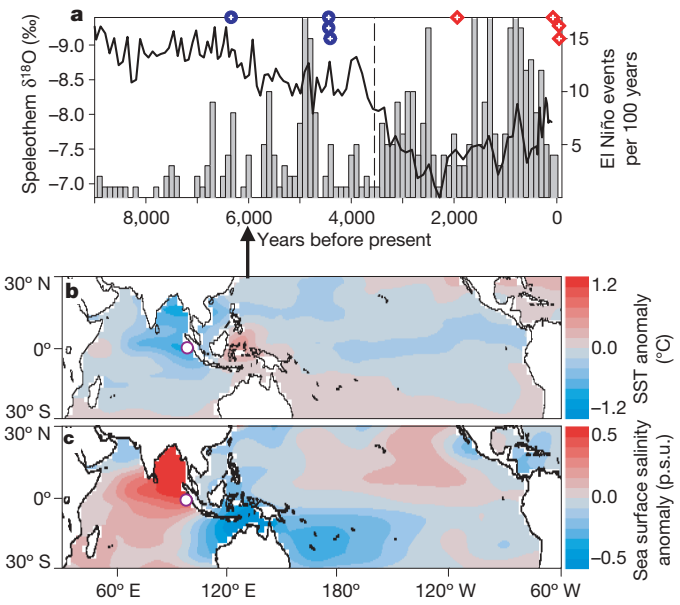


Figure 2 | Monsoon–ENSO–IOD evolution during the Holocene. **a**, Holocene evolution of the Asian summer monsoon and El Niño, as recorded by a speleothem $\delta^{18}O$ record from Dongge Cave, China[17] (black curve) and lake sediments from Laguna Pallcacocha, Ecuador[18] (grey bars), respectively. The dashed vertical line approximately marks the shift from strong monsoon activity during the early–middle Holocene to strong El Niño activity in the late Holocene. Approximate ages of IOD events reconstructed using Mentawai corals (Fig. 3 and Supplementary Fig. 3) are shown by blue circles (strong monsoon scenario) and red diamonds (strong El Niño scenario). In the eastern Indian Ocean, climate simulations (FOAM[24,25]) for 6,000 calibrated years BP produce cooler mean annual SSTs (relative to the present day) that result in a more IOD-like mean state (**b**) and a widespread increase in mean annual sea surface salinity (p.s.u., practical salinity units) (**c**). The white circles in **b** and **c** show location of the Mentawai Islands.

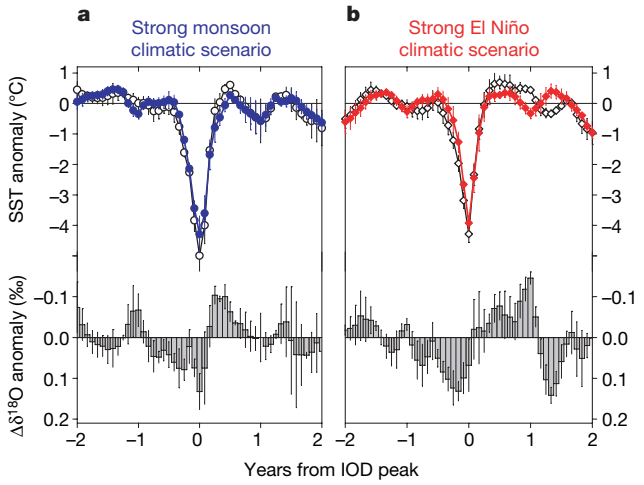


Figure 3 | Composite records of eastern IOD upwelling events. Composite coral Sr/Ca–SST (coloured curves), $\delta^{18}O$ (black curves; shown on an apparent-SST scale) and $\Delta\delta^{18}O$ (grey bars) anomalies were produced for IOD events during periods of strong monsoon activity (middle Holocene) (**a**) and strong El Niño activity (late Holocene) (**b**). Each IOD composite was produced using coral records of four individual eastern IOD upwelling events (Supplementary Fig. 3 and Supplementary Table 2). Error bars show the standard error for each monthly value of the composite IOD profiles.

©2007 Nature Publishing Group

upwelling events within both climatic scenarios implies that the monsoon wind reversal has been a consistent factor controlling the timing of peak eastern IOD cooling.

Although the magnitude and seasonal timing of peak cooling in the eastern IOD sector is similar for both climatic scenarios, the duration of the cool SST anomaly is greater during times of strong monsoon forcing. Cool anomalies exceeding 1.6 °C (that is, twice as much cooling as the average annual minimum of modern SST in the Mentawai Islands[13]) span five months under the strong monsoon forcing of the middle Holocene (Fig. 4). During the late Holocene, however, the same magnitude of IOD cooling is constrained to three months. It is unlikely that this result is an artefact related to differences in intra-annual coral growth because there is no systematic difference between the average annual cycles of Sr/Ca–SST and $\Delta\delta^{18}O$ recorded by the middle and late Holocene corals (Supplementary Fig. 4).

To examine the physical processes that may have led to extended IOD cooling during the strong monsoon scenario of the middle Holocene, we used orbitally forced simulations of Indian Ocean climate from two coupled ocean–atmosphere models, the Fast Ocean Atmosphere Model (FOAM) and NCAR-CCSM1[24,25]. Under the orbital conditions of 6,000 calibrated years BP, FOAM documents an increase in mean monsoon strength and a mean La Niña-like state in the tropical Pacific that is in good agreement with palaeoclimate reconstructions[24,25]. At the same time, the simulated mean annual SST in the Indian Ocean shows the clear signature of a positive IOD state due to cooling of the surface ocean in the eastern IOD upwelling region (Fig. 2b). This IOD-like mean state is a robust result that is obtained in experiments with both models throughout the early to middle Holocene. The simulated mean SST cooling in the eastern IOD region during the early to middle Holocene is the result of strengthened Asian monsoon winds in both the summer and winter driven by the enhanced seasonal cycle of insolation. The enhanced Asian summer monsoon produces anomalously strong southeasterly winds in the Mentawai region, while the enhanced winter monsoon generates anomalous northeasterly winds in the northeastern Indian Ocean. The easterly component of these anomalous winds produces upwelling and surface cooling in the equatorial eastern Indian Ocean throughout the year.

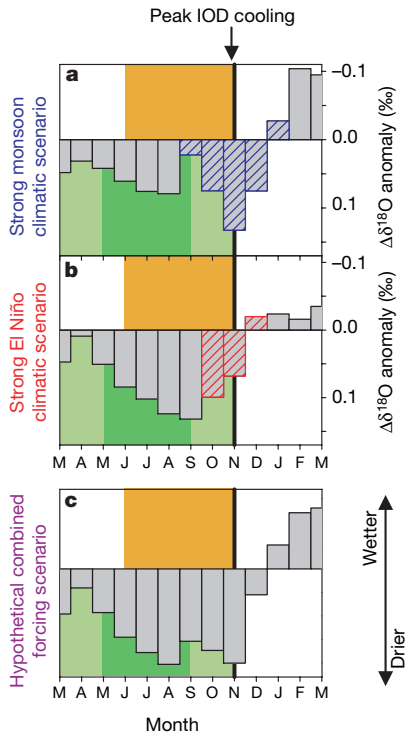


Figure 4 | Composite records of IOD drought. Composite IOD drought signals during a single monsoon cycle for the strong monsoon (**a**) and strong El Niño climatic scenarios (**b**). Bars with coloured hatching show $\Delta\delta^{18}O$ values that correspond to strong (exceeding 1.6 °C) cool SST anomalies. Vertical line denotes peak IOD cooling. Approximate periods of the transition and summer phases of the Asian monsoon are denoted by light and dark green shading, respectively. Orange shading shows the present-day timing of El Niño-driven droughts in western Indonesia. **c**, Also shown is a hypothetical 'combined IOD drought' scenario where IOD conditions are forced synergistically both by strong El Niño and by strong monsoon conditions, such as could occur with projected ENSO-independent strengthening of the Asian monsoon.

The simulation results indicate that the prolonged cooling during mid-Holocene IOD events may reflect an earlier development of anomalous Ekman upwelling in the eastern IOD sector, due to stronger cross-equatorial winds driven by enhanced summer warming of the Asian landmass in the middle Holocene[17,25]. Another possibility, suggested by the simulations of mid-Holocene monsoon dynamics, is that prolonged cooling during IOD events could result from stronger summer monsoon winds superimposed on a strengthened equatorial easterly wind field. The prolonged IOD cooling during the strong monsoon scenario is also consistent with model studies of IOD initiation processes, which suggest that IOD events triggered by processes internal to the Indian Ocean tend to begin earlier than those triggered by El Niño events in the Pacific[7,8].

The composite IOD reconstructions show that drought is a key feature associated with IOD events in both the strong monsoon and strong El Niño climatic scenarios (Fig. 3). In both scenarios the $\Delta\delta^{18}O$ anomalies suggest that rainfall is reduced for almost the entire year leading up to peak IOD conditions. However, the nature and timing of peak drought conditions are quite different for the two climatic scenarios (Fig. 4). During the strong El Niño setting of the late Holocene, the $\Delta\delta^{18}O$ record indicates that eastern IOD droughts have a single main dry phase that is synchronous with the timing of El Niño-driven droughts in western Indonesia (~June–November)[26]. Moreover, the drought peaks before strong (greater than 1.6 °C) IOD cooling occurs (Fig. 4b). Hence, it appears that the relatively brief period of strong IOD cooling in this climatic scenario restricts eastern IOD droughts primarily to the duration of drying expected from El Niño conditions alone. The composite IOD reconstruction for the strong El Niño climatic scenario also shows a period of drought in the second year after peak IOD conditions (Fig. 3). This is a robust feature of each of the individual IOD events from the late Holocene (Supplementary Fig. 3), and may be related to the biennial nature of the IOD noted in some studies[1,2].

In contrast, droughts in the eastern IOD region during the middle Holocene tended to be bimodal and peaked later (Fig. 4a). In this strong monsoon scenario, the early phase of IOD drought corresponds with the Asian summer monsoon season (~May–September) and is probably caused by strengthened cross-equatorial airflow drawing moisture from the tropical Indian Ocean into the Asian monsoon. A second, and more prominent, dry phase develops later and is synchronous with the extended period of strong (greater than 1.6 °C) IOD cooling in this climatic scenario. This suggests that prolonged cooling in the tropical eastern Indian Ocean can also play a direct role in driving IOD droughts in western Indonesia.

These findings are supported by our simulations of mid-Holocene sea surface salinity (Fig. 2c), which show a region of increased salinity extending from the coast of Sumatra into the Bay of Bengal. This salinity anomaly is associated with enhanced surface-ocean evaporation driven by strengthened Asian monsoon winds. The simulated cooling of SSTs in the tropical eastern Indian Ocean (Fig. 2b) also drives anomalous atmospheric subsidence and divergence, which reduces precipitation. These eastern IOD drought processes during the middle Holocene are also consistent with the simulated droughts produced in models of El Niño-independent IOD events[8].

The coral records and model simulations demonstrate that both ENSO and the Asian monsoon are capable of influencing drought

©2007 Nature Publishing Group

dynamics in the eastern IOD sector. Historically these influences on the IOD have tended to oppose each other, because El Niño conditions have generally led to weaker monsoons. However, there is evidence indicating that this tight coupling of the Asian monsoon and ENSO systems has begun to break down over recent decades[6]. The weakening of this relationship may be due to systematic changes in the pattern of Pacific SST anomalies during El Niño events[27], or due to transient greenhouse warming of the Eurasian landmass which has enabled ENSO-independent strengthening of the Asian monsoon[6,21,22]. During the twenty-first century it is expected that greenhouse warming will cause the Asian monsoon to strengthen further[21,22]; however, light-scattering aerosols and land-clearing may initially reduce monsoon strength before aerosol-control policies and greenhouse-gas emissions lead to an abrupt return of a more intense Asian monsoon[23].

Our findings raise the possibility that the extent of the rural challenges posed by future changes in Asian monsoon strength may have been underestimated. In particular, the later peak that we observe in IOD droughts during the strong monsoon scenario of the middle Holocene coincides with the time when rainfall in western Indonesia is normally at its annual maximum (Supplementary Fig. 4). Hence, the consequences of such a change in the timing of maximum IOD drought conditions may be particularly severe. Recent changes in the ENSO–monsoon relationship further raise the possibility that IOD droughts could intensify in the future if monsoon forcing acts synergistically with El Niño (Fig. 4c). Hence, while predictions of changing monsoon strength have focused on impacts in Asia and India[23,28], any ENSO-independent strengthening of the Asian monsoon will probably serve to prolong IOD-related droughts in western Indonesia, and possibly also southern Australia[3], thus adding to the socio-economic impacts of the IOD.

Received 10 May; accepted 21 November 2006.

1. Webster, P. J., Moore, M. D., Loschnigg, J. P. & Leben, R. R. Coupled ocean–atmosphere dynamics in the Indian Ocean during 1997–98. *Nature* **401,** 356–360 (1999).
2. Saji, H. H., Goswami, B. N., Vinayachandran, P. H. & Yamagata, T. A dipole mode in the tropical Indian Ocean. *Nature* **401,** 360–363 (1999).
3. Ashok, K., Guan, Z. & Yamagata, T. Influence of the Indian Ocean Dipole on the Australian winter rainfall. *Geophys. Res. Lett.* **30,** 1821, doi:10.1029/2003GL017926 (2003).
4. Abram, N. J., Gagan, M. K., McCulloch, M. T., Chappell, J. & Hantoro, W. S. Coral reef death during the 1997 Indian Ocean Dipole linked to Indonesian wildfires. *Science* **301,** 952–955 (2003).
5. Birkett, C., Murtugudde, R. & Allan, T. Indian Ocean climate event brings floods to East Africa's lakes and the Sudd Marsh. *Geophys. Res. Lett.* **26,** 1031–1034 (1999).
6. Kumar, K. K., Rajagopalan, B. & Cane, M. A. On the weakening relationship between the Indian Monsoon and ENSO. *Science* **284,** 2156–2159 (1999).
7. Behera, S. K. *et al.* A CGCM study on the interactions between IOD and ENSO. *J. Clim.* **19,** 1688–1705 (2006).
8. Fischer, A. S., Terray, P., Guilyardi, E., Gualdi, S. & Delecluse, P. Two independent triggers for the Indian Ocean Dipole/Zonal Mode in a coupled GCM. *J. Clim.* **18,** 3428–3449 (2005).
9. Charles, C. D., Cobb, K. M., Moore, M. D. & Fairbanks, R. G. Monsoon-tropical ocean interaction in a network of coral records spanning the 20th century. *Mar. Geol.* **201,** 207–222 (2003).
10. Beck, J. W. *et al.* Sea-surface temperatures from coral skeletal strontium/calcium ratios. *Science* **257,** 644–647 (1992).
11. Gagan, M. K. *et al.* Temperature and surface-ocean water balance of the mid-Holocene tropical western Pacific. *Science* **279,** 1014–1018 (1998).
12. Briggs, R. W. *et al.* Deformation and slip along the Sunda Megathrust in the great 2005 Nias-Simeulue earthquake. *Science* **311,** 1897–1901 (2006).
13. Reynolds, R. W. & Smith, T. M. Improved global sea surface temperature analyses using optimum interpolation. *J. Clim.* **7,** 929–948 (1994).
14. Charles, C. D., Hunter, D. E. & Fairbanks, R. G. Interaction between the ENSO and the Asian monsoon in a coral record of tropical climate. *Science* **277,** 925–928 (1997).
15. Eaton, P. & Radojevic, M. *Forest Fires and Regional Haze in Southeast Asia* (Nova Science Publishers, Huntington, New York, 2001).
16. Sontakke, N. A., Pant, G. B. & Singh, N. Construction of all-India summer monsoon rainfall series for the period 1844–1991. *J. Clim.* **6,** 1807–1811 (1993).
17. Yuan, D. *et al.* Timing, duration, and transitions of the Last Interglacial Asian monsoon. *Science* **304,** 575–578 (2004).
18. Moy, C. M., Seltzer, G. O., Rodbell, D. T. & Anderson, D. M. Variability of El Nino/Southern Oscillation activity at millennial timescales during the Holocene epoch. *Nature* **420,** 162–165 (2002).
19. Gagan, M. K., Hendy, E. J., Haberle, S. G. & Hantoro, W. S. Post-glacial evolution of the Indo-Pacific Warm Pool and El Nino-Southern Oscillation. *Quat. Int.* **118–119,** 127–143 (2004).
20. Koutavas, A., Lynch-Stieglitz, J., Marchitto, T. M. & Sachs, J. P. El Nino-like pattern in Ice Age tropical Pacific sea surface temperature. *Science* **297,** 226–230 (2002).
21. Ashrit, R. G., Kumar, K. R. & Kumar, K. K. ENSO-monsoon relationships in a greenhouse warming scenario. *Geophys. Res. Lett.* **28,** 1727–1730 (2001).
22. Hu, Z.-Z., Latif, M., Roeckner, E. & Bengtsson, L. Intensified Asian summer monsoon and its variability in a coupled model forced by increasing greenhouse gas concentrations. *Geophys. Res. Lett.* **27,** 2681–2684 (2000).
23. Zickfeld, K., Knopf, B., Petoukhov, V. & Schellnhuber, H. J. Is the Indian summer monsoon stable against global change? *Geophys. Res. Lett.* **32,** L15707, doi:10.1029/2005GL022771 (2005).
24. Liu, Z., Brady, E. & Lynch-Stieglitz, J. Global ocean response to orbital forcing in the Holocene. *Paleoceanography* **18,** 1041, doi:10.1029/2002PA000819 (2003).
25. Liu, Z., Harrison, S. P., Kutzbach, J. & Otto-Bliesner, B. Global monsoons in the mid-Holocene and oceanic feedback. *Clim. Dyn.* **22,** 157–182 (2004).
26. Dai, A. & Wigley, T. M. L. Global patterns of ENSO-induced precipitation. *Geophys. Res. Lett.* **27,** 1283–1286 (2000).
27. Kumar, K. K., Rajagopalan, B., Hoerling, M., Bates, G. & Cane, M. Unraveling the mystery of Indian monsoon failure during El Nino events. *Science* **314,** 115–119 (2006).
28. Anderson, D. M., Overpeck, J. T. & Gupta, A. K. Increase in the Asian southwest monsoon during the past four centuries. *Science* **297,** 596–599 (2002).
29. Xie, P. & Arkin, P. A. Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Clim.* **9,** 840–858 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Prayudi, I. Suprianto, K. Glenn, T. Watanabe, H. Scott-Gagan, K. Sieh and the Indonesian Institute of Sciences (LIPI) for assistance with fieldwork, which was carried out under research permits issued by LIPI. We thank H. Scott-Gagan, J. Cali, G. Mortimer, A. Alimanovic and D. Kelleher for laboratory assistance, and R. Gallimore and P. Behling for the model simulations and model output processing. This study was supported by an Australian Postgraduate Award and RSES Jaeger Scholarship to N.J.A., and an Australian Research Council Discovery grant to M.K.G. and W.S.H.

Author Contributions N.J.A. was responsible for coral geochemical analysis and interpretation of the records. M.K.G. was Chief Investigator and the Australian Institutional Counterpart for the ARC project. Z.L. provided climate model simulations and advice on ocean–atmosphere interactions. W.S.H. was Partner Investigator and the Indonesian Institutional Counterpart for the ARC project. M.T.M. supported the TIMS Sr/Ca analyses. B.W.S. provided extensive logistical support for fieldwork. N.J.A and M.K.G. wrote the paper, with comments provided by all other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to N.J.A. (nabr@bas.ac.uk) and M.K.G. (michael.gagan@anu.edu.au).

©2007 Nature Publishing Group

LETTERS

Complex gas hydrate from the Cascadia margin

Hailong Lu[1], Yu-taek Seo[1]†, Jong-won Lee[1]†, Igor Moudrakovski[1], John A. Ripmeester[1], N. Ross Chapman[2], Richard B. Coffin[3], Graeme Gardner[4] & John Pohlman[3]

Natural gas hydrates are a potential source of energy[1] and may play a role in climate change[2] and geological hazards[3]. Most natural gas hydrate appears to be in the form of 'structure I', with methane as the trapped guest molecule[4], although 'structure II' hydrate has also been identified, with guest molecules such as isobutane and propane, as well as lighter hydrocarbons[5,6]. A third hydrate structure, 'structure H', which is capable of trapping larger guest molecules, has been produced in the laboratory[7], but it has not been confirmed that it occurs in the natural environment. Here we characterize the structure, gas content and composition, and distribution of guest molecules in a complex natural hydrate sample recovered from Barkley canyon, on the northern Cascadia margin[8]. We show that the sample contains structure H hydrate, and thus provides direct evidence for the natural occurrence of this hydrate structure. The structure H hydrate is intimately associated with structure II hydrate, and the two structures contain more than 13 different hydrocarbon guest molecules. We also demonstrate that the stability field of the complex gas hydrate lies between those of structure II and structure H hydrates, indicating that this form of hydrate is more stable than structure I and may thus potentially be found in a wider pressure–temperature regime than can methane hydrate deposits.

Relatively few natural hydrate samples have been recovered sufficiently intact to have been characterized completely by a suite of laboratory techniques. Ideally, such a characterization includes determination of the structure and the composition, gas analysis and the distribution of the guests over the sites in the hydrate cages, as well as the levels of water saturation in the sediment and the degree of water conversion to hydrate. Some attempts were made towards achieving this goal with samples recovered from the Mallik 5L-38 site[9] and from a cold seep on the Cascadia margin[10]. In addition to the geological markers, only a detailed characterization of the hydrate itself can give clues as to the initial formation of hydrate as well as predict the stability fields and decomposition characteristics.

The hydrate samples characterized in this Letter were recovered from massive outcrops on the sea floor in Barkley canyon during a survey with the Canadian Remotely Operated Platform for Ocean Science (ROPOS), about 80 km off the west coast of Vancouver island[8]. This is near the site of an accidental recovery of hydrate by a commercial fishing boat that was dragging a trawl net on the canyon bottom[11]. Hydrate samples used in this work were chipped from a hydrate outcrop with the mechanical arm of ROPOS and stored in a stainless steel pressure chamber that was part of the arm: this kept the hydrate at pressure during the ascent to the ship. Once on board, the chamber was vented, the hydrate removed and stored in liquid nitrogen, the entire process taking less than one minute. Subsequently the sample was transported to the laboratory by air transport in a cold dry-shippable container.

An oil-stained sample taken from a water depth of ~855 m was sub-sampled for total gas analysis (see Methods), which showed that there were 183 volumes of gas recovered from a volume of water (at 25 °C), thus 76% of the water was in the hydrate phase if it is assumed that all cages in the hydrate were filled, more if some of the cages were empty. The gas components (Supplementary Fig. 1) present at levels above 0.10 mol% of the gas released from the dissociated hydrate are listed in Table 1. They are predominantly structure II (sII) hydrate formers, but 10.6 mol% of the hydrocarbons with C ≥ 2 are potential structure H (sH) hydrate guests suitable for the large cage, including C5, C6 and C7. Surprisingly, this also includes *n*-C5 and *n*-C6, for which it was not possible to make a sH hydrate in the laboratory[12]. Although the presence of sH hydrate can be inferred from the gas composition, a direct measurement of hydrate structure and a confirmation of guests in hydrate cages is necessary, as some previously unknown hydrate formers were detected.

The powder X-ray diffraction pattern obtained at 85 K is shown in Fig. 1. After being calibrated with ice I_h reflections, a common contaminant in hydrate samples, and on comparing the pattern with standard hydrate patterns, the main reflections can be indexed as those of sII, consistent with the analytical results on gas compositions. However, a number of weak reflections are not characteristic of either of structure I (sI) or sII hydrate, but will fit the pattern of sH hydrate. The structural characteristics of the sample studied are consistent with the presence of both cubic sII and hexagonal sH hydrate (Supplementary Table 1), with the following parameters; $a =$

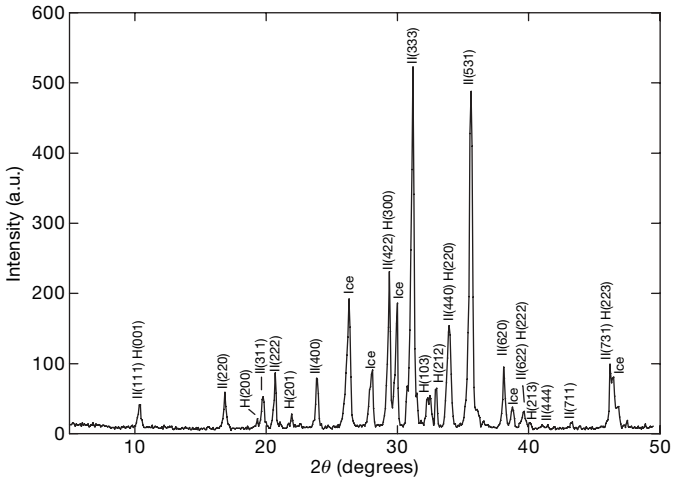


Figure 1 | The powder X-ray diffraction pattern of the gas hydrate studied, from Barkley canyon, offshore Vancouver island. The prefix to each reflection label indicates the hydrate: I, hydrate I; II, hydrate II; H, hydrate H. a.u., arbitrary units.

[1]Steacie Institute for Molecular Sciences, National Research Council of Canada Ottawa, Ontario, Canada K1A 0R6. [2]School of Earth and Ocean Sciences, University of Victoria, Victoria, Canada V8W 3P6. [3]Naval Research Laboratory, Washington DC 20375, USA. [4]Institute of National Measurement Standards, National Research Council Canada, Ontario, Canada K1A 0R6. †Present addresses: New Energy Research Department, Korea Institute of Energy Research, 71-2, Daejeon 305-343, Korea (Y.S.); Department of Environmental Engineering, Konju National University, Chungnam 330-717, Korea (J.L.).

©2007 Nature Publishing Group

Table 1 | Composition of recovered natural gas and NMR data on synthesized hydrates

Gas composition		Content (mol%)	Location* and corresponding NMR chemical shift				
			sII		sH		
			5^{12}	$5^{12}6^4$	5^{12}	$4^35^66^3$	$5^{12}6^8$
C1	Methane	75.36	−4.3	−8.2	−4.5	−4.9	
C2	Ethane	11.41		6.0			
C3	Propane	5.66		16.7, 17.5			
C4	Isobutane	1.22		26.3, 23.5			
C4	butane	2.16		13.7, 25.8			
C5	Isopentane	0.71					11.6, 22.4, 30.7, 32.4
C5	Pentane	0.34					13.7, 21.6, 34.5
C6	2-methylpentane	0.19					12.9, 19.4, 22.9, 25.1
C6	3-methylpentane	0.15					11.8, 17.6, 30.2, 36.9
C6	*n*-hexane	0.22					13.2, 20.8, 28.1
C6	Methylcyclopentane	0.44					20.3, 25.6, 35.4, 35.9
C6	Cyclohexane	0.54		27.7			
C7	Methylcyclohexane	0.73		0.53			23.6, 27.3, 33.7, 36.1
Benzene		0.53					-

Shown are the enclathrated gas components in the cages of gas hydrates recovered from Barkley canyon, offshore Vancouver island, and the corresponding ^{13}C NMR chemical shifts (p.p.m.) determined on synthetic gas hydrates.
* Cage in hydrate structure. Nomenclature: $5^{12}6^4$, for example, indicates that the cage is constructed of 12 pentagons and 4 hexagons.

1.724 ± 0.002 nm and $a = 1.219 \pm 0.009$ nm, $c = 1.006 \pm 0.009$ nm, respectively.

A ^{13}C NMR spectrum obtained on another sub-sample was quite complex, with a number of sharp resonances and broader overlapping signals that cover a range from −8 p.p.m. to ~36 p.p.m. On comparing the chemical shifts in Fig. 2 with those of the synthetic gas hydrates listed in Table 1, the sharp signals are characteristic of CH_4 in the large cage of sII and the small cages of sH and sII, and ethane, propane and isobutane in the large cage of sII. Other hydrocarbons suitable for the large cage in sII are *n*-butane and cyclohexane, both of which require the presence of a small cage guest to give stable hydrates; benzene, also a sII guest, was observed in the gas chromatography-mass spectrometry, but we did not examine the aromatic spectral region by NMR. The remaining signals are interpreted as belonging to isopentane, methylcyclopentane and methylcyclohexane in the large sH cage, consistent with their relatively high abundance among sH hydrate formers as indicated by the gas chromatographic results. Table 1 also indicates that hydrocarbons trapped in hydrate cages usually have unique shifts that are different from the pure hydrocarbon phases (Supplementary Table 2), an important factor in identifying guest species in hydrates. Indeed, the chemical shifts characteristic of *n*-pentane and *n*-hexane in the large sH cage were confirmed from work recently carried out to illustrate the ability of these materials to act as guests for sH when they are present in combination with other hydrate formers such as neohexane[13]. Although the sH hydrate formers are present in relatively small quantities, it must be remembered that each large-cage guest requires from four to five methane molecules to give a stable structure.

Figure 3 shows pressure–temperature (*P*–*T*) curves for a number of sII, sH and sI hydrates. The sII double hydrates containing methane are the most stable over the temperature range investigated. The plot for the mixed sII-H hydrate falls between those for the sII double hydrates and sH hydrate. This suggests that sH in the mixed sII-H

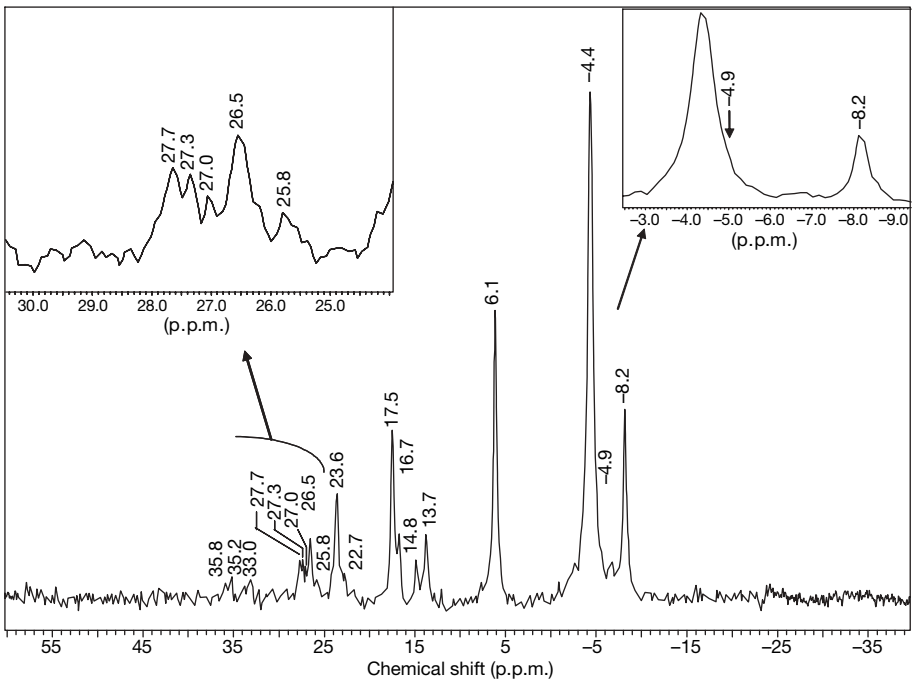


Figure 2 | Solid state ^{13}C NMR spectra of the gas hydrate from Barkley Canyon. Assignments are given here as chemical shift (in p.p.m.), gas name, cage name, and hydrate type. −8.2, methane, $5^{12}6^4$, sII; −4.9, methane, $4^35^66^3$, sH; −4.4, methane, 5^{12}, sII and sH; 6.1, ethane, $5^{12}6^4$, sII; 16.7 and 17.5, propane, $5^{12}6^4$, sII; 13.7 and 25.8, *n*-butane, $5^{12}6^4$, sII; 23.6 and 26.5, isobutane, $5^{12}6^4$, sII; 27.7, cyclohexane, $5^{12}6^4$, sII; 25.8, 35.15 and 35.8 p.p.m., methylcyclopentane, $5^{12}6^8$, sH; 23.6 and 27.3, methylcyclohexane, $5^{12}6^8$, sH; 22.7 and 33.0, isopentane, $5^{12}6^8$, sH.

©2007 Nature Publishing Group

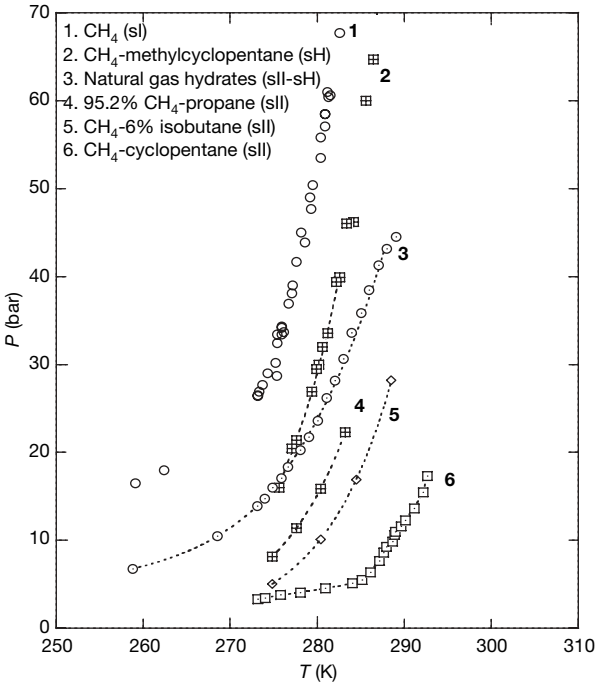


Figure 3 | The dissociation conditions for natural gas hydrates from Barkley canyon and reference samples of sI, sII and sH gas hydrates.

sample decomposes first, following more or less the curve for the pure sH hydrate. After that, the *P*–*T* curve becomes parallel to the sII hydrate curves. All of the methane-containing double hydrates are considerably more stable than sI methane hydrate. It is clear that the complex hydrate will have a much greater regime of stability in natural environments than sI methane hydrate. From the evidence presented, the complex gas hydrate is composed of sII and a smaller amount of sH hydrate. The two structures appear to be independent but closely associated. Although Sassen and MacDonald[6] and Pohlman *et al.*[14], on the basis of gas composition measurements, conjectured that sH hydrate existed in the Gulf of Mexico and Barkley canyon, offshore Vancouver island, respectively, they did not analyse the hydrate structures directly, nor were the heavy hydrocarbon gas components confirmed to be present in the hydrate cages. Thus our research provides the first confirmation of the existence of natural sH hydrate after its natural existence was predicted in 1987[7].

Because the complex hydrate is much more stable than methane hydrate, the region where this hydrate may be found is greatly expanded over that of methane hydrate, possibly appearing at water depths as shallow as about 200 m, and at a depth of 500 m below the sea floor in the Barkley canyon region (Supplementary Fig. 2). As the stability region of complex hydrates has increased by a factor of ~5 in *P*–*T* space, increasing the depth range of hydrate stability in the natural environment by a factor of ~3, it is clear that a substantial occurrence of such complex gas hydrates must be expected likely in previously unsuspected locations, such as shallow water or below the base of the gas hydrate stability zone for methane hydrate.

Hydrocarbon gases originate from two sources, biogenic and thermogenic, the main indicator being the $\delta^{13}C$ value with respect to the PDB standard, $\delta^{13}C_{PDB}$, of the methane in the hydrocarbons[15]: biogenic < -60‰, thermogenic > -60‰. The $\delta^{13}C_{PDB}$ of methane in the complex hydrate was found to be -46‰, and the fractional abundance of methane in the hydrocarbons, $C_1/\Sigma C_{2+}$, was 3.1, indicative of the thermogenic origin of enclathrated hydrocarbons. Gas hydrates of thermogenic hydrocarbons have been reported in many sites—for example, the Gulf of Mexico[6,16], the Caspian Sea[17], the Black Sea[18,19] and the Japan Sea[20]. Because most thermogenic hydrocarbons are generated at temperatures above 100 °C, about 1,000–2,000 m (or even deeper) below the sea floor in most areas and well below the stability zone of methane hydrate, it is more likely that these will be found in deep sediment. Those found nearer the sea floor are associated with fluid or gas seepage. Because of this, complex thermogenic gas hydrate associated with seepage will occur near the sea floor, as observed in the Gulf of Mexico and northern Cascadia, but also in sediments below the base of methane hydrate stability zone. The appearance of C_4 and C_5 in the lower parts of the methane hydrate stability zone has been confirmed at Blake ridge[21], and it has been proposed[22,23] (on the basis of the results of gas composition analyses) that the deeper of the double BSRs (bottom simulating reflectors) identified on Hydrate ridge, offshore Oregon, might have been caused by the formation of sII and sH gas hydrates. With the drilling of deeper wells, it is likely that more thermogenic complex gas hydrates will be discovered. On the other hand, complex gas hydrates occurring at shallow water depths will be far more subject to the influence of sea level changes than those located in deep water. One can well imagine massive dissociation of complex gas hydrate present at shallow depths if the sea level were to drop by some 150 m, an event that happened during the last glaciation period[24].

METHODS

Sample handling. For the characterization, sample handling was always in liquid nitrogen, with sub-samples removed for total gas analysis, gas composition analysis, powder X-ray diffraction, ^{13}C nuclear magnetic resonance (NMR) spectroscopic analysis and stability field measurements.

Total gas analysis. A portion (~ 200 mg) of gas hydrate sample was pumped down on a vacuum line at a sample temperature of 120 K to remove adsorbed nitrogen[25]. The amount of gas released from hydrate samples can be calculated from the gas pressure build-up in a known volume[26].

Gas composition determination. A piece (~100 mg) of gas hydrate was placed in a specially designed sample vessel, allowing evacuation and gas sampling through a rubber sealing cap. The gas compositions were identified with a Varian Saturn 2200 mass spectrometer, then quantitatively determined with a Varian CP 3800 gas chromatograph.

Powder X-ray diffraction analysis. This was carried out on a Rigaku θ–θ powder X-ray diffractometer at 85 K (radiation was CoK_α, wavelength 1.79021 Å). About 100 mg of powdered sample was loaded into a flat bed sample holder set in a PAAR TTK-HC temperature control system. The temperature was adjusted by varying the liquid nitrogen flow through the assembly, and was monitored with a platinum resistance thermometer placed directly beneath the sample holder. The instrumental conditions for the experiments were set as: current 25 mA, voltage 45 V and a scan range in 2θ of 5–50°. The patterns obtained were calibrated with several ice I_h reflections.

^{13}C solid-state NMR measurements. These were carried out on a Bruker DSX-400 NMR spectrometer at a frequency of 100.6 MHz. Finely powdered samples (about 200 mg) were cold-loaded into ZrO_2 spinners and spectra were obtained at 173 K in order to avoid hydrate decomposition. The chemical shift region between -40 and $+60$ p.p.m. (as referred to tetramethylsilane, TMS) was recorded, which covers the region of most hydrocarbons that can stabilize hydrate structures as guest materials. Spectra were obtained under conditions deemed to give reliable spectra for quantification, with a ^{13}C pulse length of 4.7 μs (90°), high power ^1H decoupling, and delay times of 30 s. Usually 250–500 scans were accumulated in order to obtain a good signal to noise ratio.

Determination of hydrate stability. In order to establish the stability zone of the mixed sII-H hydrate, a sub-sample (~3 g) was placed in a small-volume pressure vessel at low temperature and evacuated to remove air. The temperature was then allowed to increase in a stepwise fashion and kept constant after each temperature rise until the pressure in the system became constant.

Received 31 July; accepted 17 November 2006.

1. Kleinberg, R. & Brewer, P. Probing gas hydrate deposits — Exploiting this immense unconventional energy resource presents great challenges. *Am. Sci.* **89**, 244–251 (2001).
2. Xu, W., Lowell, R. P. & Peltzer, E. T. Effect of seafloor temperature and pressure variations on methane flux from a gas hydrate layer: Comparison between current and late Paleocene climate conditions. *J. Geophys. Res.* **106**, 26413–26424 (2001).
3. Dillon, W. P. *et al.* in *Natural Gas Hydrates: Occurrence, Distribution, and Dynamics* (eds Paull, C. K. & Dillon, W.P.) 211–233 (AGU Monogr. 124, American Geophysical Union, Washington DC, 2001).
4. Ginsburg, G. D. & Soloviev, V. A. *Submarine Gas Hydrates* (Norma Publishers, St Petersburg, Russia, 1998).
5. Davidson, D. W. *et al.* Laboratory analysis of a naturally occurring gas hydrate. *Geochim. Cosmochim. Acta* **50**, 619–623 (1986).

©2007 Nature Publishing Group

6. Sassen, S. & McDonald, I. R. Evidence of structure H hydrate, Gulf of Mexico continental slope. *Org. Geochem.* **22,** 1029–1032 (1994).
7. Ripmeester, J. A., Tse, J. S., Ratcliffe, C. I. & Powell, B. M. A new clathrate hydrate structure. *Nature* **325,** 135–136 (1987).
8. Chapman, N. R., Pohlman, J., Coffin, R., Chanton, J. & Lapham, L. Thermogenic gas hydrates in the Northern Cascadia Margin. *Eos* **85,** 361–368 (2004).
9. Ripmeester, J. A. *et al.* Structure and composition of gas hydrate in sediment recovered from the JAPEX/JNOC/GSC et al. Mallik 5L–38 gas hydrate production research well, determined by X-ray diffraction and Raman and solid-state nuclear magnetic resonance spectroscopy. *Geol. Surv. Can. Bull.* **585,** 106 (2005).
10. Lu, H. *et al.* The occurrence and structural characteristics of the gas hydrates associated with a cold vent field, Offshore Vancouver Island. *J. Geophys. Res.* **110,** B10204, doi:10.1029/2005JB003900 (2005).
11. Spence, G. D., Chapman, N. R., Hyndman, R. D. & Cleary, C. Fishing trawler nets massive "catch" of methane hydrates. *Eos* **82,** 621–627 (2001).
12. Ripmeester, J. A. & Ratcliffe, C. I. ^{129}Xe NMR studies of clathrate hydrates: new guests for structure II and structure H hydrates. *J. Phys. Chem.* **94,** 8773–8776 (1990).
13. Lee, J., Lu, H., Moudrakovski, I. L., Ratcliffe, C. I. & Ripmeester, J. A. n-pentane and n-hexane as coguests in a structure H hydrate in mixtures of 2,2-dimethylbutane and methane. *Angew. Chem. Int. Edn Engl.* **45,** doi:10.1002/anie.200504366 (2006).
14. Pohlman, J. W. *et al.* The origin of the thermogenic gas hydrates on the northern Cascadia Margin as inferred from isotopic (^{13}C/^{12}C and D/H) and molecular composition of hydrate and vent gas. *Org. Geochem.* **36,** 703–716 (2005).
15. Bernard, B. B., Brooks, J. M. & Sackett, W. M. Natural gas seepage in the Gulf of Mexico. *Earth Planet. Sci. Lett.* **31,** 48–58 (1976).
16. Sassen, R. *et al.* Free hydrocarbon gas, gas hydrate, and authigenic minerals in chemosynthetic communities of the northern Gulf of Mexico continental slope: relation to microbial processes. *Chem. Geol.* **205,** 195–217 (2004).
17. Diaconescu, C. C., Kieckhefer, R. M. & Knapp, J. H. Geophysical evidence for gas hydrates in the deep water of the South Caspian Basin, Azerbaijan. *Mar. Petrol. Geol.* **18,** 209–221 (2001).
18. Woodside, J. M., Modin, D. I. & Ivanov, M. K. An enigmatic reflector on subbottom profiler record from the Black Sea — the top of the shallow gas hydrate deposits. *Geo-Mar. Lett.* **23,** 269–277 (2003).
19. Mazzini, A. *et al.* Methane-related authigenic carbonates from the Black Sea: geochemical characterization and relation to seeping fluid. *Mar. Geol.* **212,** 153–181 (2004).
20. Matsumoto, R. *et al.* Gas hydrate layer and prominent flares of gas plumes in Naoetsu Basin, Eastern margin of Japan Sea. *Eos Trans.* **86**(AGU Fall meet. suppl.), abstr. OS41C-04 (2005).
21. Kvenvolden, K. A. & Bernard, L. A. Gas hydrates of the Blake Outer Ridge, Site 533, Deep Sea Drilling Project Leg 76. *Init. Rep. DSDP* **76,** 353–365 (1983).
22. Milkov, A. V. *et al.* Ethane enrichment and propane depletion in subsurface gases indicate gas hydrate occurrence in marine sediments at southern Hydrate Ridge offshore Oregon. *Org. Geochem.* **35,** 1067–1080 (2004).
23. Milkov, A. V., Claypool, G. E., Lee, Y.-J. & Sassen, R. Gas hydrate systems at Hydrate Ridge offshore Oregon inferred from molecular and isotopic properties of hydrate-bound and void gases. *Geochim. Cosmochim. Acta* **69,** 1007–1026 (2005).
24. Lambeck, K., Esat, T. M. & Potter, E.-K. Links between climate and sea levels for the past three million years. *Nature* **419,** 199–206 (2002).
25. Tulk, C. A., Ratcliffe, C. I. & Ripmeester, J. A. Chemical and physical analysis of natural gas hydrate from the JAPEX/JNOC/GSC Mallik 2L–38 gas hydrate research well. *Geol. Surv. Can. Bull.* **544,** 263–267 (1999).
26. Lu, H., Dutrisac, R., Ripmeester, J. A., Wright, F. & Uchida, T. Measurement of gas hydrate saturation in sediment cores recovered from the JAPEX/JNOC/GSC et al. Mallik 5L–38 gas hydrate production research well. *Geol. Surv. Can. Bull.* **585,** 89 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.A.R. (john.ripmeester@nrc-cnrc.gc.ca).

©2007 Nature Publishing Group

LETTERS

A critical ligamentous mechanism in the evolution of avian flight

David B. Baier[1], Stephen M. Gatesy[1] & Farish A. Jenkins Jr[2]

Despite recent advances in aerodynamic[1,2], neuromuscular[3–5] and kinematic[6,7] aspects of avian flight and dozens of relevant fossil discoveries[8], the origin of aerial locomotion and the transition from limbs to wings continue to be debated[9,10]. Interpreting this transition depends on understanding the mechanical interplay of forces in living birds, particularly at the shoulder where most wing motion takes place. Shoulder function depends on a balance of forces from muscles, ligaments and articular cartilages, as well as inertial, gravitational and aerodynamic loads on the wing[11]. Here we show that the force balance system of the shoulder evolved from a primarily muscular mechanism to one in which the acrocoracohumeral ligament has a critical role. Features of the shoulder of Mesozoic birds and closely related theropod dinosaurs indicate that the evolution of flight preceded the acquisition of the ligament-based force balance system and that some basal birds are intermediate in shoulder morphology.

Steady-speed gliding, in comparison to flapping, is more amenable to estimating the interplay of forces stabilizing the shoulder in flight. The minimum forces needed to hold the wing outstretched in a gliding pigeon are provided by a mathematical model[12]. Each wing produces a vertical aerodynamic force at its centre of pressure[13] of approximately 0.5BW (0.5 × body weight; Fig. 1a, b). This force engenders a moment about the shoulder (tending to elevate the wing relative to the body) that is countered by a downward force from the pectoralis muscle. The pectoralis acts with such a short lever arm (L_p, Fig. 1b) that a force of 7BW is required to balance the aerodynamic moment.

However, this model[12] inaccurately represents the shoulder as a sturdy ball-and-socket joint capable of fully supporting the wing. In reality, a large, ovoid humeral head articulates with a relatively shallow, saddle-shaped glenoid with open dorsal and ventral margins[14] (Fig. 1c). The glenoid is structurally incapable of resisting the downward pull on the humerus as the pectoralis counters the elevation moment. Antagonistic muscles and/or ligaments must provide an upward force to preclude ventral dislocation.

The primary upstroke muscle, the supracoracoideus, might be expected to pull upward on the humerus, but its tendon actually pulls ventromedially (rotating the wing by an elevation/supination moment[15] rather than by raising the humeral head). Only three of the twelve muscles crossing the shoulder are oriented to produce an upward force (scapulohumeralis, latissimus dorsi and the subcoracoscapulares complex; Table 1, F_{ymax}) in gliding. The acrocoracohumeral ligament (AHL, Fig. 1c), spanning from the acrocoracoid process of the coracoid to the transverse sulcus of the humerus, is also situated to preclude ventral dislocation. The ligament was predicted to limit wing pronation (in pigeon[16]), or constrain the normal range of flapping movements (in starling[14]). In swimming penguins, the AHL was inferred to be taut in all wing positions, effectively coordinating shoulder movement and preventing disarticulation[17]. To explore the relative contributions of these four structures to joint stability, we undertook a force balance analysis of a wing in a gliding posture.

Three-dimensional force balance calculations that include the AHL yield a family of viable solutions in which all forces and moments equilibrate, requiring ligament forces of about 6–9BW and only small inputs from other muscles (Table 1; Fig. 2a, b). No combination of the three upward-pulling muscles can balance the system without the AHL. Higher forces (about 13BW) are generated by the pectoralis during flapping[18], implying even greater demands on the AHL. Mechanical testing of the pigeon AHL reveals that the ligament can withstand tensile loading of 39BW (± 5BW, $n = 10$) without failure, far exceeding potential flight-induced forces. Following transection of the AHL, stimulation of the pectoralis or manipulation of the humerus in fresh specimens produces ventral shoulder dislocation, as expected. We conclude that the AHL is critical to maintaining the integrity of the shoulder joint during aerial locomotion. Rather than being supported from below by articular cartilage, the humeral head is suspended from above by the AHL.

The avian coracoid is typically described as a compressive strut between the wing and the sternum[12,13] that resists loads generated by the pectoralis. Although the pectoralis vector generally aligns with the long axis of the coracoid, the laterally facing glenoid cannot by itself transmit a force from the humeral head down the coracoid's shaft. At the glenoid, our force balance predicts that the coracoid bears a medially directed load equal and opposite to the laterally directed joint force on the humeral head (F_j; Fig. 2a, b). If unopposed, this force would push the coracoids together, compressing the thorax. The medial push on the glenoid is primarily balanced by the lateral pull of the AHL. The acrocoracoid process experiences a ventrolateral force equal and opposite to the dorsomedially directed AHL force experienced by the proximal humerus (F_{ahl}; Fig. 2a, b). Compression at the glenoid and tension on the acrocoracoid process sum to a net ventromedial force that approximately aligns with the shaft of the coracoid. This net force can be primarily opposed by a dorsolateral force from the sternum at the coracosternal joint. The AHL thus provides a mechanism for transmitting the pectoralis force through the coracoid as a compressive strut.

Has AHL function evolved as a specialization for flight in birds or is its role primitive for archosaurs? Crocodilians, the closest extant relatives of birds, possess a homologous coracohumeral ligament (CHL) and, like birds, retain the ancestral archosaurian saddle-shaped glenoid[14]. However, in alligators, as in most other diapsids, the long axis of the glenoid is oriented approximately 90° to that of birds, such that the open margins of the joint are anterior and posterior (Fig. 2c). The alligator's humerus is supported ventrally by the coracoid portion of the glenoid saddle, but during locomotion

[1]Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island 02912, USA. [2]Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA.

©2007 Nature Publishing Group

Table 1 | Force magnitudes and force balance solutions

	F_{max}	F_{ymax}	S1	S2	S3	S4
M. pectoralis	22.4	−19.7	5.9	5.1	6.1	6.0
M. supracoracoideus	10.2	−3.8	0	0	0	0
M. deltoideus, pars minor	1.1	−0.4	0	0	0.3	0.6
M. deltoideus, pars major	0.7	−0.1	0	0	0	0.4
M. scapulohumeralis	5.5	**3.2**	1.2	0	1.2	0.8
M. coracobrachialis cranialis	0.7	−0.2	0	0	0.3	0
M. coracobrachialis caudalis	3.1	−2.1	0	0	0	0
M. latissimus dorsi	0.5	**0.2**	0	0	0	0
M. biceps brachii	4.0	−0.6	0.2	0.2	0	0
M. scapulotriceps	4.5	−0.4	0	0	0	0
M. subcoracoscapulares	6.0	**2.9**	0	2.2	0	0.8
M. deltoideus, pars propatagialis	2.7	−0.3	0.5	0.6	0	0
Lig. acrocoracohumerale	39.0	**20.7**	8	6.0	8.5	8.1
Joint contact force	-	-	8.5	8.5	8.8	9.5
Total muscle force	-	-	7.8	8.1	7.9	8.6

Maximum force (F_{max}) for muscles and ligaments crossing the shoulder joint of a pigeon was derived from physiological cross sectional area measurements, mechanical testing and published data. Maximum vertical force (F_{ymax}) is only upward (positive, bold) for three muscles and the AHL. Columns S1–S4 are four of the many possible 3D force balance solutions for gliding (details in Methods). All values are in multiples of body weight.

retractors such as the caudal part of the pectoralis pull the humeral head towards the open, posterior margin.

Unlike birds, crocodilians possess several muscles that arise from a broad scapulocoracoid plate anterior to the glenoid (Fig. 3a) and are thus capable of countering the pectoralis. Three-dimensional kinematic records of the shoulder girdle and humerus of alligators walking on the treadmill were compared with ligament-taut positions derived from manipulating dissections. Minimal overlap between ligament-taut and walking humeral positions demonstrates that the CHL is slack throughout the majority of the stride cycle and therefore not essential for joint stabilization during locomotion. Electromyographic data from varanid lizards[19] show that several muscles, despite their protractive lines of action, are active while the limb is retracting, suggesting a stabilizing role. Homologous protractor muscles in *Alligator mississippiensis* are also likely to be primarily responsible for balancing the pectoralis. We propose that shoulder joint stabilization evolved within archosaurs from a primarily active, muscle-based balance system to a passive, ligament-based system in extant, volant birds.

On the basis of comparative coracoid morphology in *Deinonychus*, *Archaeopteryx* and extant birds, the coracoid tuberosity of theropod dinosaurs is considered to be homologous with the avian acrocoracoid process[20]. With reference to the tuberosity/process and other anatomical landmarks, we reconstructed the orientation of the scapulocoracoid and CHL/AHL for the non-avialan theropods *Sinornithoides youngi* and *Sinornithosaurus millenii* and for the volant avialans *Archaeopteryx lithographica* and *Confuciusornis sanctus* (Fig. 3b–e). The ancestral ligament orientation of the CHL is retained in all but *C. sanctus* (Fig. 3e), in which elevation of the acrocoracoid process above the glenoid and deepening of the humeral head yield an

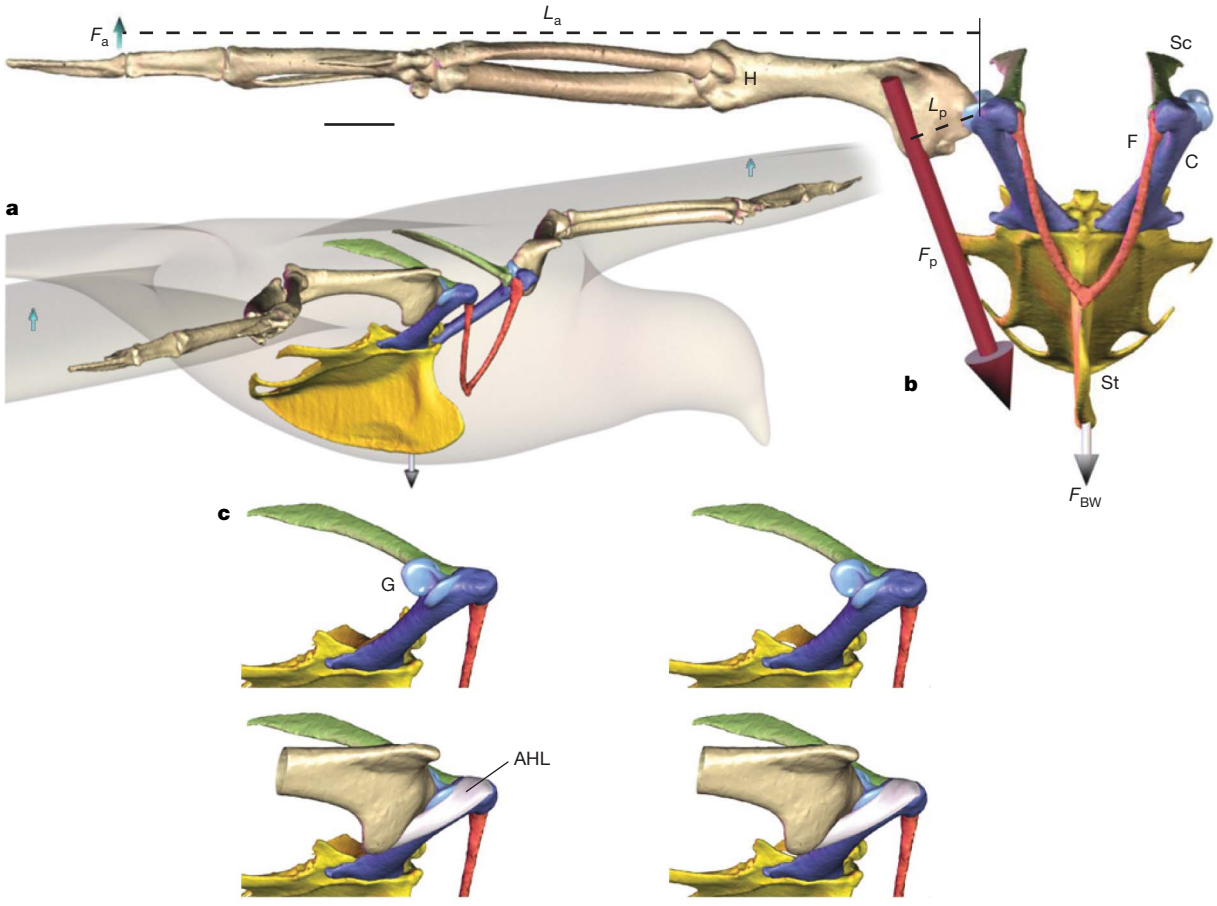


Figure 1 | Forelimb skeleton and pectoral girdle of a pigeon. a, **b**, Anterolateral (**a**) and anterior (**b**) views. During a steady speed glide the bird's body weight (downward grey vector shown below the chest, F_{BW}) is supported by aerodynamic forces of 0.5BW from each wing (upward blue vectors, F_a). F_a has a lever arm (L_a) about the shoulder approximately 11× that of the pectoralis (L_p) requiring a pectoralis force (red vector, F_p) of about 6BW to balance the elevation moment. **c**, Anterolateral stereo pairs showing the saddle shaped glenoid (light blue, G), which lacks a ventral shelf to support the humeral head. The acrocoracohumeral ligament (AHL), spanning from the elevated acrocoracoid process of the coracoid to the transverse sulcus on the proximal humerus, is ideally situated to prevent ventral dislocation by the pectoralis. H, humerus, Sc, scapula (green), C, coracoid (blue), F, furcula (red), St, sternum (yellow). Scale bar in **b** equals 1 cm.

©2007 Nature Publishing Group

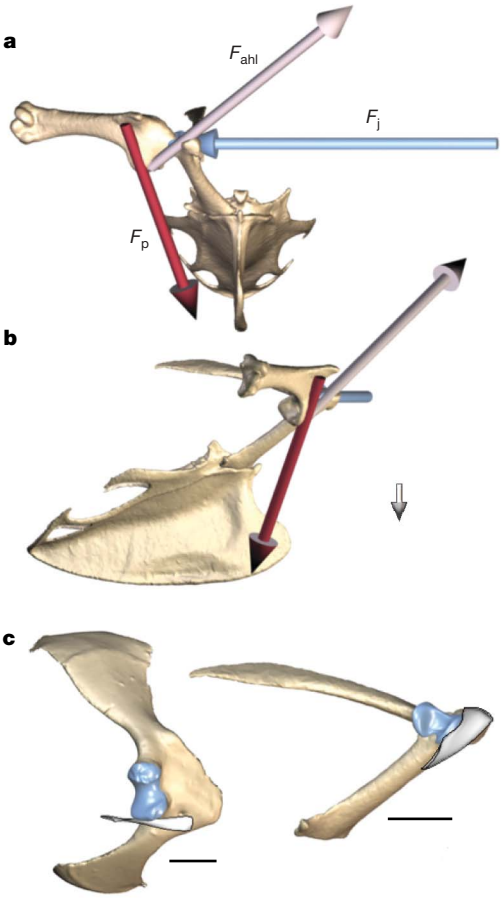


Figure 2 | Force balance for the glenohumeral joint of a gliding pigeon and comparison of alligator and pigeon shoulder morphology. a, **b**, Anterior (**a**) and right lateral (**b**) views of primary forces balancing the shoulder (S1, Table 1). The ventromedial force of the pectoralis (F_p) is resisted by the dorsomedial pull of the acrocoracohumeral ligament (F_{ahl}) and a lateral push from the glenoid (F_j). Aerodynamic and smaller muscle forces are not shown. Grey vector proportionately indicates one body weight. **c**, Comparison of the alligator (left) and pigeon (right) right scapulocoracoids in lateral view, showing the saddle-shaped glenoid cartilages (blue) and homologous coracohumeral and acrocoracohumeral ligaments (white, cut at humeral attachment). Scale bars equal 1 cm.

AHL orientation intermediate between that in crocodilians (Fig. 3a) and extant birds (Fig. 3f).

An intermediate condition is also found in the recently discovered *Jeholornis*[21] and *Sapeornis*[22], which diverged from the line to extant birds between *Archaeopteryx* and *Confuciusornis*[22]. Hence, we conclude that the common ancestor of *Jeholornis* and extant birds had likely developed an incipient stabilizing role of the AHL (Fig. 3). On the basis of coracoid morphology, full development of the ligament-based force balance traces back at least to the common ancestor of Enantiornithes and Aves[23].

The selective impetus for transforming the coracoid tuberosity into an acrocoracoid process has been associated with the functional modification of three muscles. Displacement of the acrocoracoid directly affected the lines of action of the coracobrachialis and biceps brachii (coracoid head) because, like the AHL, both originate from the process[20]. In addition, enlargement of the tuberosity formed the triosseal canal, creating a pulley that deflected the tendon of the supracoracoideus such that this muscle became the primary humeral elevator/supinator powering upstroke[15]. We now also recognize elevation and expansion of the acrocoracoid process as directly related to the reorientation and increased loading of the AHL, which underlies the novel osseo-ligamentous mechanism for passive shoulder stabilization in modern birds.

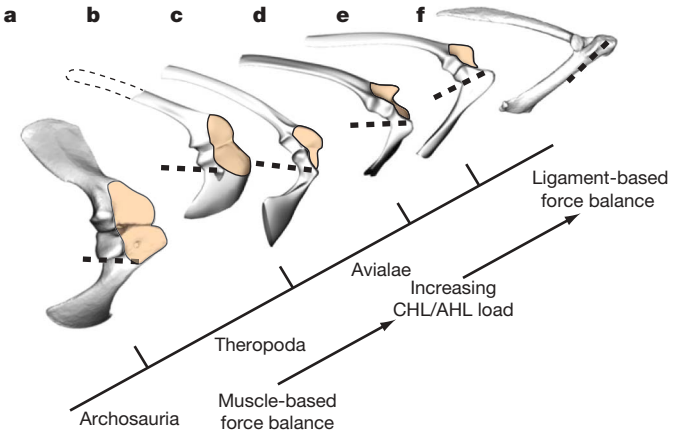


Figure 3 | Comparison of archosaur scapulocoracoids. Lateral view of right scapulocoracoids of *Alligator mississippiensis* (**a**), *Sinornithoides youngi* (**b**), *Sinornithosaurus millenii* (**c**), *Archaeopteryx lithographica* (**d**), *Confuciusornis sanctus* (**e**) and *Columba livia* (**f**) on a simplified phylogeny[23]. The changing orientation of the CHL/AHL (dashed lines) and the reduction of the scapulocoracoid plate anterior to the glenoid for muscle attachment (orange surfaces) record the transition to a novel, ligament-based force balance mechanism in modern birds. The dashed lines do not represent ligament length.

METHODS

Pigeon model. Bones of a wild-caught pigeon (*Columba livia*) were scanned at the High-Resolution X-ray Computed Tomography Facility at the University of Texas at Austin (UTCT). Three-dimensional polygonal models were created using VTK (Kitware), articulated into a virtual skeleton in Maya (Autodesk), and configured in a mid-downstroke posture (humeral shaft protracted 60° from the vertebral axis, elevated 0° and pronated 15° relative to a neutral position with distal condyles vertical). On the basis of dissections, muscular and ligamentous attachments were mapped onto the virtual skeleton to determine their positions and lines of action (Fig. 1).

Three-dimensional force balance. The wing was treated as a free rigid body articulating only at the shoulder (elbow and wrist joints fixed). Aerodynamic, gravitational, muscular, joint and ligament forces were mapped onto the virtual skeleton as 16 vectors in Maya. These 3D vectors were used to calculate three translational components (F_x, F_y, F_z) and three rotational moments (M_x, M_y, M_z) relative to the coordinate system's origin at the point of glenohumeral contact (Fig. 2). Two vectors are fully described. A net aerodynamic force (F_a) of 0.5BW acts vertically upward at the centre of pressure, which was estimated to be 13.2 cm lateral and 1.1 cm posterior to the origin by averaging the quarter-chord points of ten chord-wise strips of equal area along the wing's 29.3 cm shoulder–wingtip span[13]. The wing's weight of 0.074BW acts vertically downward at its centre of mass (3.2 cm lateral and 0.8 cm posterior to the origin, based on a frozen, outstretched wing). All other vectors are only partially characterized. Muscle and ligament lines of action are fixed by anatomy and wing posture in all but the pectoralis, for which a range of vector directions were tested owing to its broad origin. Likewise, joint contact force orientation is constrained to horizontal by glenoid morphology. Maximum isometric muscle force (F_{max}) was estimated from the physiological cross sectional area of one pigeon and from published data[24–26] (Table 1). However, the actual magnitudes of muscle, joint and ligament forces are unknown, creating an indeterminate system for which multiple answers exist.

A family of potential solutions was explored by interactively solving the 3D force balance using a MEL (Maya embedded language)-based expression to identify equilibrium. We discovered viable combinations of muscle activation using a simple workflow. All muscle forces began completely inactive (0BW), but could be increased up to F_{max}. First, a pectoralis force (F_p) was added to balance the aerodynamic elevation moment (wing weight was included but was not significant). Second, horizontal forces created by the pectoralis were balanced by the joint force (F_j). Third, the pectoralis's vertical force was countered—primarily by the ligament force (F_{ahl}) because muscles alone proved insufficient. Other muscles were then interactively recruited to balance remaining forces and moments, automatically adjusting previously entered F_p, F_j and F_{ahl} values as required. We sought simple alternatives that did not invoke unnecessary and widespread co-contraction (Fig. 2a, b; Table 1). The necessity for high AHL loading was insensitive to variation in pectoralis orientation or wing posture.

©2007 Nature Publishing Group

Mechanical testing. The coracoid–AHL–humerus complex was dissected from frozen pigeon specimens, each bone potted in low melting-temperature metal alloy and then mounted on an Instron Mechanical Testing System (Model #4442) with a 500 N load cell. Tensile force was measured for 15 loading/unloading cycles from 0–70 N (5 ligaments from 3 birds), followed by failure tests (10 ligaments from 6 birds). Crosshead speeds (rate of loading) ranged between 25–1000 mm min^{-1}, with data collected at either 6.67 or 20 points s^{-1}. The ligament substance was not completely torn in any of the failure tests. Failure occurred by humeral (6) or acrocoracoid (4) fracture.

Alligator shoulder data. Three alligators (1.08–3.08 kg) were recorded at 30 frames per second with simultaneous light (Sony DCR-TRV30) and dorsoventral fluoroscopic video (Sony DCR-VX1000 digital handycam attached to a 27.94 cm image intensifier with a Siemens cineradiographic apparatus) as they walked on a motorized treadmill. Optical and radiographic distortions were removed by 'warping' images of square calibration grids in Maya. Shoulder and forelimb bone models were created using a 3D laser scanner (ShapeGrabber), articulated, and animated in Maya by aligning each bone to its fluoroscopic image[27]. Shoulder kinematic data during walking were compared with ligament-taut positions derived by matching models to digital photographs of manipulated specimens.

Fossil reconstructions. Three-dimensional models of the scapula, coracoid and humerus were created in Maya based on measurements, photographs and drawings of the original material as well as published descriptions[14,20,28–30]. The vertebral column, sternum, furcula and ribs (when present) were used to constrain the position and orientation of the scapulocoracoids. The thoracic vertebral column was oriented horizontally, with the humerus in a position comparable to that of the gliding pigeon model.

Received 21 July; accepted 10 November 2006.
Published online 17 December 2006.

1. Hedrick, T. L., Usherwood, J. R. & Biewener, A. A. Wing inertia and whole-body accelerations: an analysis of instantaneous aerodynamic force production in cockatiels (*Numphicus hollandicus*) flying across a range of speeds. *J. Exp. Biol.* **207,** 1689–1702 (2004).
2. Spedding, G. R., Rosén, M. & Hedenström, A. R. A family of vortex wakes generated by a thrush nightingale in free flight in a wind tunnel over its entire natural range of flight speeds. *J. Exp. Biol.* **206,** 2313–2344 (2003).
3. Biewener, A. A. & Dial, K. P. *In vivo* strain in the humerus of pigeons (*Columba livia*) during flight. *J. Morphol.* **255,** 61–75 (1995).
4. Jenkins, F. A., Dial, K. P. & Goslow, G. E. J. A cineradiographic analysis of bird flight: the wishbone in starlings is a spring. *Science* **241,** 1495–1498 (1988).
5. Goslow, G. E. J., Wilson, D. & Poore, S. O. Neuromuscular correlates to the evolution of flapping flight in birds. *Brain Behav. Evol.* **55,** 85–99 (2000).
6. Tobalske, B. W. Neuromuscular control and kinematics of intermittent flight in the European starling (*Sturnus vulgaris*). *J. Exp. Biol.* **198,** 1259–1273 (1995).
7. Tobalske, B. W., Peacock, W. L. & Dial, K. P. Kinematics of flap-bounding flight in the zebra finch over a wide range of speeds. *J. Exp. Biol.* **202,** 1725–1739 (1999).
8. Chiappe, L. M. & Witmer, L. M. *Mesozoic Birds* (Univ. California Press, Berkeley, 2002).
9. Dial, K. P. Wing-assisted incline running and the evolution of flight. *Science* **299,** 402–404 (2003).
10. Xu, X. *et al.* Four-winged dinosaur from China. *Nature* **421,** 335–340 (2003).
11. Gatesy, S. M. & Baier, D. B. The origin of the avian flight stroke: a kinematic and kinetic perspective. *Paleobiology* **31,** 382–399 (2005).
12. Gray, S. J. *Animal Locomotion.* (Weidenfeld and Nicolson, London, 1968).
13. Pennycuick, C. J. The strength of the pigeon's wing bones in relation to their function. *J. Exp. Biol.* **46,** 219–233 (1967).
14. Jenkins, F. A. The evolution of the avian shoulder joint. *Am. J. Sci.* **293,** 253–267 (1993).
15. Poore, S. O., Sánchez-Haiman, A. & Goslow, G. E. J. Wing upstroke and the evolution of flapping flight. *Nature* **387,** 798–802 (1997).
16. Sy, M. Funktionell-anatomische untersuchungen am vogelflügel. *J. Ornithol.* **84,** 199–296 (1936).
17. Bannasch, R. Morphologisch-funktionelle untersuchung am lokomotionsapparat der pinguine als grundlage fur ein allgemeines bewegungsmodell des unterwasserfluges. *Gegenbaurs Morphol. Jahrb.* **132,** 757–817 (1986).
18. Usherwood, J. R., Hedrick, T. L., McGowan, C. P. & Biewener, A. A. Dynamic pressure maps for wings and tails of pigeons in slow, flapping flight, and their energetic implications. *J. Exp. Biol.* **208,** 355–369 (2005).
19. Jenkins, F. A. & Goslow, G. E. J. The functional anatomy of the shoulder of the Savannah monitor lizard (*Varanus exanthematicus*). *J. Morphol.* **175,** 195–216 (1983).
20. Ostrom, J. H. Some hypothetical stages in the evolution of avian flight. *Smithson. Contrib. Paleobiol.* **27,** 1–21 (1976).
21. Zhou, Z. & Zhang, F. A long-tailed, seed-eating bird from the Early Cretaceous of China. *Nature* **418,** 405–409 (2002).
22. Zhou, Z. & Zhang, F. Anatomy of the primitive bird *Sapeornis chaoyangensis* from the early cretaceous of Liaoning, China. *Can. J. Earth Sci.* **40,** 731–747 (2003).
23. Norell, M. A. & Clarke, J. A. Fossil that fills a critical gap in avian evolution. *Nature* **409,** 181–184 (2001).
24. Dial, K. P. & Biewener, A. A. Pectoralis muscle force and power output during different modes of flight in pigeons (*Columba livia*). *J. Exp. Biol.* **176,** 31–54 (1993).
25. Poore, S. O., Ashcroft, A., Sánchez-Haiman, A. & Goslow, G. E. J. The contractile properties of the M. supracoracoideus in the pigeon and starling: a case for long-axis rotation of the humerus. *J. Exp. Biol.* **200,** 2987–3002 (1997).
26. Woolley, J. D. The functional morphology of the avian flight muscle M. coracobrachialis posterior. *J. Exp. Biol.* **203,** 1767–1776 (2000).
27. Gatesy, S. M., Dial, K. P. & Jenkins, F. A. An inside look at skeletal motion in flying birds. *J. Vert. Paleo.* **24,** 63A (2004).
28. De Beer, G. Archaeopteryx lithographica*: A Study Based upon the British Museum Specimen* (The Trustees of the British Museum, London, 1954).
29. Paul, G. S. *Dinosaurs of the Air* (The Johns Hopkins University Press, Baltimore, 2002).
30. Chiappe, L. M., Ji, S., Ji, Q. & Norell, M. A. Anatomy and systematics of the Confuciousornithidae (Theropoda: Aves) from the Late Mesozoic of Northeastern China. *Bull. Am. Mus. Nat. Hist.* **242,** 1–89 (1999).

Acknowledgements We are grateful to K. Dial for the pigeon specimen, K. Middleton for help with VTK, C. Sullivan and L. Claessens for sharing alligator video and specimens, Z. Zhou for access to fossils at IVPP, T. Roberts for methodological advice, and the Brown Morphology group. We thank Autodesk for Maya software support and the SHAPE lab at Brown University for access to 3D laser scanning equipment. Funding was provided by the National Science Foundation (S.M.G., SHAPE lab), a Bushnell Faculty Research Grant (S.M.G.), the Paleontological Society (D.B.B.) and Sigma Xi (D.B.B.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.B.B. (David_Baier@brown.edu).

©2007 Nature Publishing Group

LETTERS

Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning

Randall D. Dahn[1], Marcus C. Davis[1], William N. Pappano[2] & Neil H. Shubin[1,3]

The genetic mechanisms regulating tetrapod limb development are well characterized, but how they were assembled during evolution and their function in basal vertebrates is poorly understood. Initial studies report that chondrichthyans, the most primitive extant vertebrates with paired appendages, differ from ray-finned fish and tetrapods in having *Sonic hedgehog* (*Shh*)-independent patterning of the appendage skeleton[1]. Here we demonstrate that chondrichthyans share patterns of appendage *Shh* expression, *Shh* appendage-specific regulatory DNA, and *Shh* function with ray-finned fish and tetrapods[2–10]. These studies demonstrate that some aspects of *Shh* function are deeply conserved in vertebrate phylogeny, but also highlight how the evolution of *Shh* regulation may underlie major morphological changes during appendage evolution.

As the most primitive extant vertebrates with paired appendages, chondrichthyans (cartilaginous fish) hold a special place in our understanding of evolution. Accordingly, analyses of the molecular pathways regulating chondrichthyan fin development should clarify our understanding of how the genetic circuitry patterning tetrapod limbs evolved. Initial studies on chondrichthyans led to a paradox: one of the fundamental signalling pathways controlling tetrapod limb development seems to have no role in patterning chondrichthyan fins[1]. In both tetrapod and teleosts, *Sonic hedgehog* (*Shh*) is a major upstream factor in development, regulating limb outgrowth and anteroposterior patterning of the skeleton[2–5]. In both taxa, *Shh* expression is restricted to the posterior mesoderm, and in tetrapods specifies both the number of skeletal elements formed and their identities[2–5]. Exogenous retinoic acid can induce ectopic anterior *Shh*, which causes mirror-image skeletal duplications in tetrapods[2,4–6]. In addition, orthologous DNA regulatory elements controlling appendage-specific *Shh* expression are common to teleost fish and tetrapods[7–10]. Chondrichthyans, however, reportedly lack both fin bud *Shh* expression and appendage-specific regulatory elements[1,9], implying *Shh*-dependent anteroposterior patterning was an innovation of bony fish (osteichthyans). It was thus proposed that *Shh* primitively functioned to free paired fins from the body wall in the common ancestor of ray-finned fish and lobe-finned fish (including tetrapods)[1]. Subsequently, *Shh* acquired its role in patterning the appendage anteroposterior axis.

We re-examined these findings by first screening chondrichthyan genomes for *Shh* appendage-specific regulatory elements (ShAREs). Remarkably, orthologous elements were found in all chondrichthyan species analysed, including chimaeras, sharks, skates and rays (Supplementary Fig. 1). Chondrichthyan ShAREs exhibit high sequence conservation with orthologues of ray-finned fish and tetrapods, and were found to occupy analogous genomic positions in intron 5 of the *Lmbr1* gene[8]. The retention of chondrichthyan ShARE sequence and genomic position strongly imply functional conservation in regulating *Shh* expression during fin development; accordingly, *Shh* orthologues from skate (*Raja erinacea*) and shark (*Chiloscyllium punctatum*) were isolated and fin bud expression assessed (phylogenetic sequence analysis presented in Supplementary Data).

As expected from the enhancer analysis, *Shh* was expressed in the posterior mesoderm of chondrichthyan fin buds, but important spatial and temporal differences with tetrapod limb buds were observed. Consistent with reported data[1], skate paired fin buds do not express *Shh* during early post-budding stages (stages 25–28, as per ref. 11; data not shown), whereas robust expression is detected in floorplate, ventrolateral neural tube, notochord and somites (Fig. 1k). At stage 29/30, subsequent to those previously examined[1], prominent

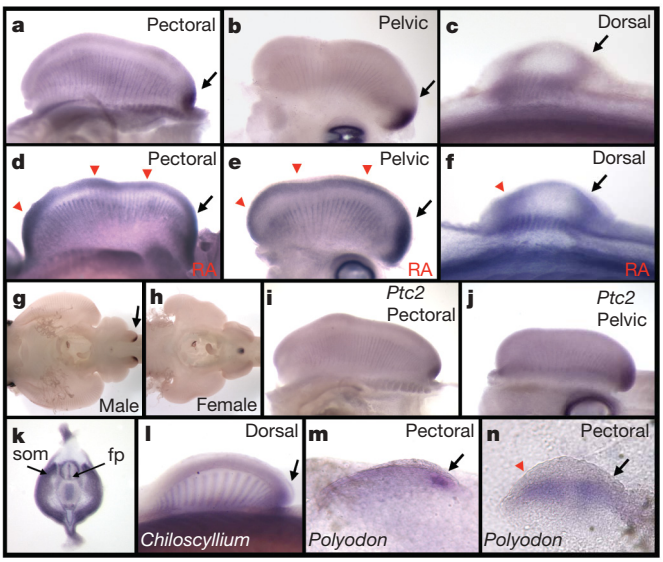


Figure 1 | *Shh* is expressed in wild-type and retinoic acid (RA)-treated basal vertebrate appendages. a–c, Stage 29/30 wild-type skate buds express *Shh* in posterior mesoderm (black arrows; *in situ* hybridizations), as in tetrapods; weak expression is also seen in distal mesoderm and muscle buds. Skate dorsal fins additionally express *Shh* anteriorly (**c**), but *Chiloscyllium* shark dorsal fins lack anterior expression (**l**). **d–f**, Retinoic-acid-treated fin buds upregulate *Shh* expression, especially in distal and ectopic anterior domains (red arrowheads). **g**, **h**, By stage 32, *Shh* appendage expression is restricted to the posterior mesoderm of claspers developing in male pelvic fins. **i**, **j**, At stage 29/30, the *Shh* target gene *Ptc2* is expressed co-extensively with Shh. **k**, Transverse sections of stage 27 skates reveal *Shh* expression in somites (som), floorplate (fp) and dorsal fin mesoderm. Parellel fin bud *Shh* expression in wild-type (**m**) and retinoic-acid-treated (**n**) *Polyodon* embryos supports a deep phylogenetic origin for *Shh* function in appendage patterning. Anterior is to the left in all panels; fin type denoted in the upper right corner.

[1]Department of Organismal Biology and Anatomy, The University of Chicago, Chicago, Illinois 60637, USA. [2]Howard Hughes Medical Institute, Department of Molecular Biology and Genetics, The Johns Hopkins School of Medicine, Baltimore, Maryland 21205, USA. [3]Field Museum of Natural History, Chicago, Illinois 60605, USA.

©2007 Nature Publishing Group

posterior border *Shh* expression was detected in both sets of paired appendages (Fig. 1a, b), but expression was never detected in anterior mesoderm. Additionally, *Shh* was weakly expressed in distal mesoderm subjacent to the apical ectoderm, and strong expression was seen in the proximodistally oriented muscle bundles along the entire anteroposterior axis (Fig. 1a, b). Posterior and distal *Shh* expression is extinguished at late stage 30 in pectoral fins, but sexually dimorphic expression of *Shh* is maintained in males along the posterior pelvic fin border for an additional 2–3 weeks (through stage 32) in the developing intermittent clasper organs (Fig. 1g, h). Expression of the *Shh* target gene *Ptc2* (ref. 12) directly mirrors the spatial and temporal domains of *Shh* expression in paired fin buds (Fig. 1i, j); skate *Ihh* was co-expressed with *Shh* in muscle buds, but asymmetric expression was never detected in posterior or anterior mesoderm (data not shown). The unpaired dorsal fin buds similarly express *Shh* in dorsal and distal mesoderm, but curiously, *Shh* transcripts were also detected in anterior mesoderm (Fig. 1c). However, we note that anterior *Shh* expression in dorsal fin buds is probably a derived condition in skates, as shark dorsal fin buds express *Shh* exclusively in posterior and dorsal mesoderm, and muscle buds (Fig. 1l).

We next demonstrated that retinoic acid regulation of *Shh* expression is conserved in chondrichthyan fin buds. *In ovo* retinoic acid treatment of stage 25–27 embryos always resulted in altered *Shh* expression in skate paired fin buds: the normal posterior and distal expression of *Shh* was strongly upregulated, ectopic anterior expression was observed, and *Shh* expression was inappropriately maintained into stage 31 (Fig. 1d, e). *Shh* spatial domains were unaffected in dorsal fin buds, but expression was upregulated (Fig. 1f). We found that retinoic-acid-mediated regulation of *Shh* expression is also conserved in the basal ray-finned fish *Polyodon spathula*, with ectopic *Shh* expression detectable in the anterior fin bud mesoderm of retinoic-acid-treated embryos in addition to the normal posterior domain (Fig. 1m, n). Combined with previous studies in zebrafish[3–5], our data indicate that regulation of *Shh* expression by retinoic acid is probably the primitive condition for all vertebrate paired appendages.

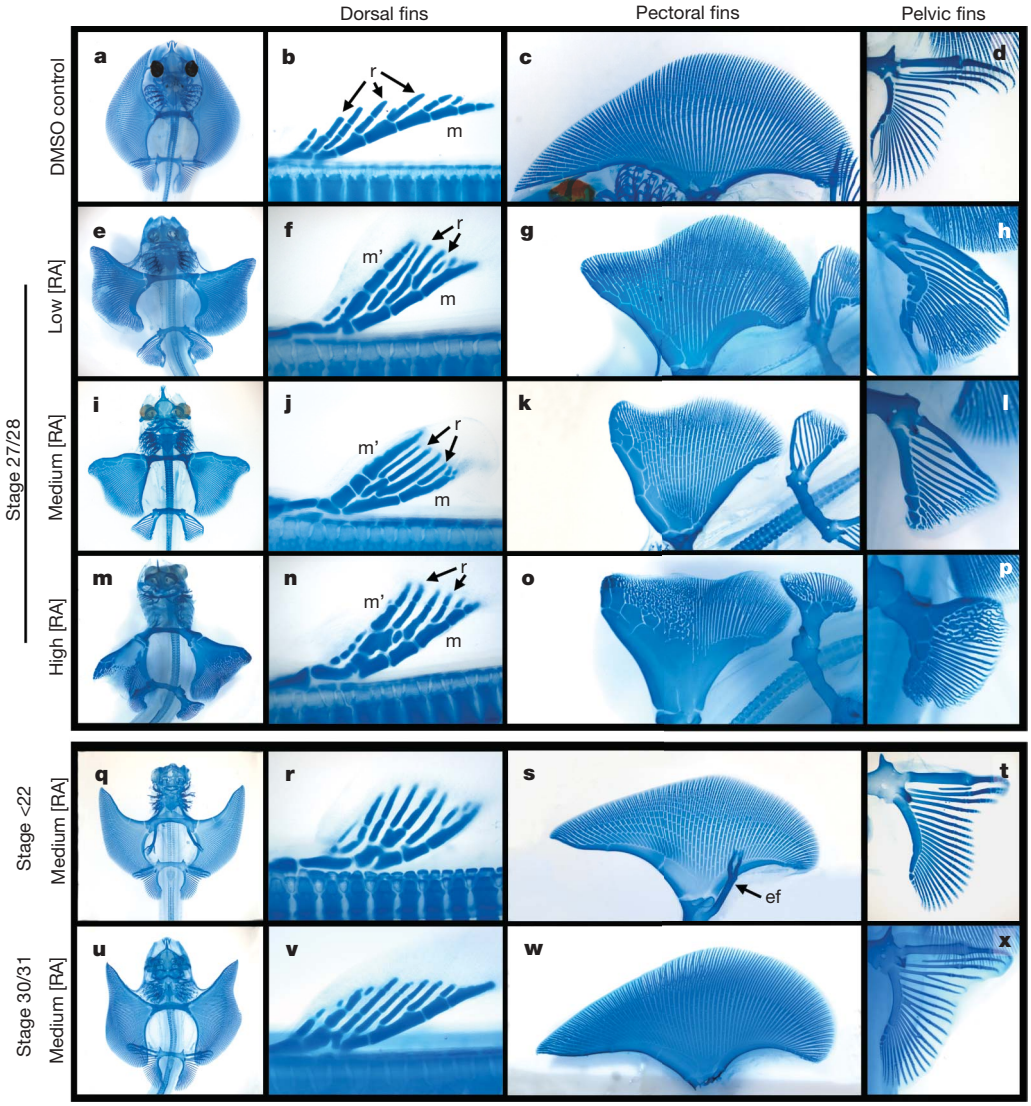


Figure 2 | Retinoic acid effects dose- and stage-dependent fin skeleton alterations in chondrichthyans. Alcian-blue-stained skeletal preparations of ~12-week-old wild-type (**a**–**d**) and retinoic-acid-treated (**e**–**x**) skates. Retinoic acid dose (low, medium and high concentrations were 0.5, 1.0, and 2.0×10^{-6} M, respectively) and stage of treatment is indicated on the left, and fin type on the top. Whole animals (**a**, **e**, **i**, **m**, **q**, **u**) and pelvic fins (**d**, **h**, **l**, **p**, **t**, **x**) are oriented with anterior to the top, whereas dorsal (**b**, **f**, **j**, **n**, **r**, **v**) and pectoral (**c**, **g**, **k**, **o**, **s**, **w**) fins are oriented with anterior to the left. Increasing retinoic acid dose exacerbates loss of anterior structures in paired fins (compare, for instance, **e**, **i**, **m**), and transforms radials into increasingly large cartilage plates (compare **g**, **k**, **o**). **q**–**t**, Early treatment induces ectopic pectoral fins (ef in **s**), whereas late treatment has mild effects (**u**–**x**). Analogous to tetrapod limbs, retinoic-acid-treated dorsal fins develop mirror-image duplications of the posterior metapterygium (m; m' for duplicated structures), and medially branching radials (r), at all doses (compare **b** with **f**, **j**, **n**, **r**, **v**).

©2007 Nature Publishing Group

Retinoic acid alters skate appendage skeletal patterns in a manner consistent with ectopic *Shh* expression, a classic phenotypic response not previously demonstrated in any non-tetrapod (Fig. 2). Most striking were the tetrapod-like mirror-image duplications observed in dorsal fins treated at stage 27/28; an ectopic metapterygium (the normal posterior-most element) forms along the anterior—as well as the posterior—fin border with radials branching medially from both (Fig. 2b, f, j, n). In pectoral fins, retinoic acid causes the deletion of anterior regions that can encompass up to half the fin, and posteriorizes the identity of skeletal elements forming along the affected anterior fin margin (Fig. 2g, k, o). Specifically, the iteratively segmented, linear rod comprising the normal anterior boundary of the fin skeleton (propterygium) is transformed to a series of bifurcating cartilaginous blocks, typical of normal metapterygial morphology

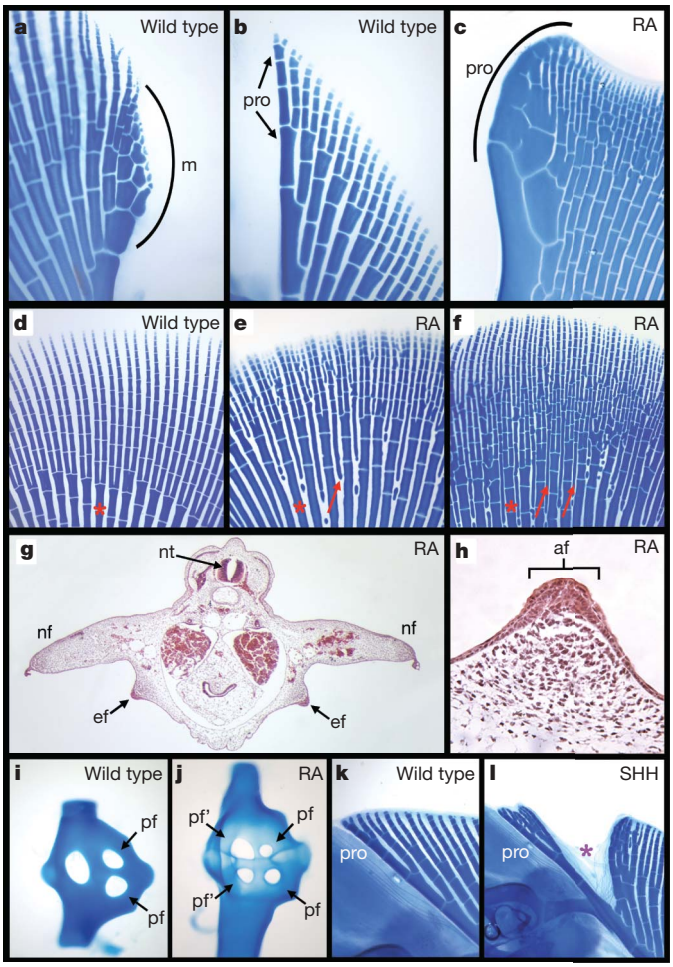


Figure 3 | Retinoic acid induces ectopic elements and posteriorizes skate fin structures. In skate pectoral fins, morphology of the posterior metapterygium (m; **a**) contrasts with the interatively segmented propterygium (pro; **b**) along the anterior. **c**, Retinoic acid treatment induces propterygia to adopt this branched pattern. Frequency of ectopic radial formation and branching also increases with retinoic acid concentration (**e**, low and **f**, high concentration versus **d**, wild type); red asterisks and arrows indicate normal and ectopic radials, respectively. Pre-stage-22 retinoic acid induces ectopic pectoral fin buds (ef; **g**) capped with pseudostratified apical ectodermal folds (af; **h**). **g**, **h** Transverse sections through stage 30 trunks with normal pectoral fins (nf) and neural tube (nt) indicated. **i**, Pectoral fin girdles normally develop with a single anterior and paired posterior (pf) foramina perforating the glenoid; retinoic acid treatment results in the anterior closely resembling posterior morphology (**j**; lateral views shown in **i** and **j**). **k**, **l**, Exogenous SHH (**l**) induces supernumerary radial formation and repolarization of cartilage branching patterns away from the bead (purple asterisk), reminiscent of the tetrapod zone of polarizing activity.

along the posterior border (compare Fig. 3a–c). Retinoic-acid-treated pelvic fins exhibit similarly transformed morphologies (Fig. 2d, h, l, p).

The ability of retinoic acid to exert dose- and stage-dependent effects on skeletal pattern[6] is also conserved in skates. Increasing doses cause progressively more severe loss of anterior regions, and more pronounced posterization of propterygial cartilages (Fig. 2a–p). Early treatment at prefin stages (stage <22) induces ectopic pectoral fin formation ventral to normal pectoral fins, as has been reported in zebrafish[13], and posteriorizes anterior pectoral girdle morphology (Fig. 2q, s; Fig. 3g–j). Later stage (stage 30/31) treatment results in a decreasingly severe suite of defects, suggesting the progressive determination of skeletal pattern during the stages examined (Fig. 2u–x). Finally, increasing retinoic acid doses increased the number of ectopic radials formed in the inter-radial spaces of paired fins (Fig. 3d–f), correlating with *Shh* upregulation in distal mesoderm (Fig. 1d, e).

Our finding that *Shh* is expressed by skate fin posterior mesoderm immediately suggests a functional parallel with tetrapod limb buds, where *Shh* expression initiates concomitant with budding and controls both skeletal element formation and identity across the entire anteroposterior axis[2,14,15]. However, the delayed onset of skate *Shh* expression (7–10 days post-budding) argues against a role in globally regulating fin anteroposterior polarity, but suggests a later role in skeletal patterning restricted to posterior elements. The physical dimensions of stage 29/30 skate fin buds support this hypothesis; *Shh* expression is restricted to the posterior tenth of the $>3{,}000\ \mu m$ long pectoral fin buds (several times the width of tetrapod limb buds), making it highly improbable that SHH protein can sufficiently diffuse to regulate patterning of any but the most proximate forming skeletal elements (see ref. 16).

The skeletal alterations induced in the anterior of skate pectoral fins by stage-31-implantation of SHH-soaked beads further support this hypothesis. Foremost, SHH effects were spatially restricted, affecting only a few radials on both sides of the bead (Fig. 3k, l). SHH re-specified the polarity of flanking radials, redirecting all cartilage branching away from the SHH source, and thus mimicking the properties that gave the tetrapod zone of polarizing activity its moniker[17]. Lastly, SHH induced formation of supernumerary radials in the flanking inter-radial spaces, which is consistent with the previous correlation between ectopic radial formation and retinoic-acid-mediated ectopic *Shh* expression (compare Fig. 1d, e to Fig. 3d–f). Thus, in combination with our previous data, we see conserved and novel roles for *Shh* signalling in patterning chondrichthyan fin buds, both of which are consistent with skates' phylogenetic position and unique fin morphologies.

Our study strongly supports the notion that a fundamental mechanism controlling anteroposterior appendage patterning, involving regulated *Shh* signalling, is deeply conserved in gnathostomes (see Supplementary Fig. 3). We have also shown that differences in the timing and spatial domains of *Shh* expression in skate fins relative to tetrapod limbs correlate well with the morphological diversification of these structures. Therefore we suggest that regulatory changes in *Shh* signalling may provide an important potential source for generating morphological diversity, in addition to the evolution of downstream *Shh*-responsive gene networks. Chondrichthyans should be an informative component of future comparative analyses of appendage development and evolution, and as the most basal extant limbed vertebrates, will be especially important in resolving the primitive genetic circuitry that was assembled to pattern the earliest fins.

METHODS

Animal husbandry and manipulation. *Raja erinacea* and *Chiloscyllium punctatum* embryos were raised at 19 °C in reconstituted Instant Ocean (Aquarium Systems), with a 12:12 h light:dark cycle. Embryos receiving AG-1X2 beads (BioRad) soaked in 1 mg ml^{-1} all-*trans* retinoic acid (Sigma) were removed from their egg cases in Instant Ocean, anesthetized with MS222 (0.4 g per 2.4 l; Sigma)

©2007 Nature Publishing Group

at room temperature for 10 min, and returned to their tanks immediately following bead implantation. Disrupting the gelatinous inner membrane of egg cases during early fin development was found to invariably result in death[11] (and data not shown), precluding the targeted implantation of retinoic-acid-soaked beads as an experimental strategy (as is employed in tetrapod embryos[6]). Alternatively, 0.33 mg ml^{-1} all-*trans* retinoic acid in DMSO was injected directly into egg cases at different volumes and stages of development, to give the final concentration indicated ($0.5–2.0 \times 10^{-6}$). Similar 'retinoic acid immersion' protocols induce ectopic anterior *Shh* expression in zebrafish pectoral fin buds[4,5], comparable to the effects of retinoic acid bead implantation studies in tetrapods[2,6].

In situ hybridization analyses were performed as previously described[1]. Alcian-blue-stained skeletal cartilage preparations, and histological sections of *Raja* embryos were generated as described[18].

Polyodon embryos were raised in 25% Hank's Balanced Salt Solution (HBSS) at 22 °C. For retinoic acid treatment, stage 37 embryos[19] were transferred to 25% HBSS containing 0.5% DMSO and 1.0×10^{-6} all-*trans* retinoic acid, incubated for 30 min in the dark, and returned to 25% HBSS. *In situ* hybridization analyses were performed as previously described[20].

Heparin acrylic beads were soaked in 1 mg ml^{-1} mouse SHH-N (gift of Phil Beachy) as previously described[21], and were implanted using the same techniques as for retinoic acid.

DNA isolation. All complementary DNAs and genomic DNA fragments were isolated using the Expand High Fidelity PCR System (Invitrogen). DNA oligonucleotide sequence can be found in Supplementary Data. GenBank accession numbers for DNA sequences are as follows: (ShARE orthologues) *R. erinacea*, EF100665; *Dasyatis sabina*, EF100655; *C. punctatum*, EF100654; *Scyliorhinus torazame*, EF100668; *Hydrolagus colliei*, EF100656; *Polypterus senegalus*, EF100660; *P. spathula* A and B, EF100661 and EF100662, respectively; *Lepisosteus osseus*, EF100658; *Amia calva*, EF100652; and turkey, EF100669; *R. erinacea Shh*, *Ihh*, *Ptc2* and *Lmbr1* intron 5, EF100667, EF100666, EF100663, and EF100657, respectively; *Raja eglanteria Lmbr1*, EF100664; *C. punctatum Shh*, EF100653; and *P. spathula Shh*, EF100659.

Sequence analysis. Molecular phylogenies were generated by aligning DNA sequences using ClustalX, and neighbour-joining trees drawn using 1,000 bootstrap trials. Phylogenetic trees were drawn using TreeViewPPC software.

Received 21 April; accepted 9 November 2006.
Published online 24 December 2006.

1. Tanaka, M. *et al.* Fin development in a cartilaginous fish and the origin of vertebrate limbs. *Nature* **416,** 527–531 (2002).
2. Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Sonic hedgehog* mediates the polarizing activity of the ZPA. *Cell* **75,** 1401–1416 (1993).
3. Krauss, S., Concordet, J. P. & Ingham, P. W. A functionally conserved homolog of the *Drosophila* segment polarity gene *hh* is expressed in tissues with polarizing activity in zebrafish embryos. *Cell* **75,** 1431–1444 (1993).
4. Akimenko, M. A. & Ekker, M. Anterior duplication of the *Sonic hedgehog* expression pattern in the pectoral fin buds of zebrafish treated with retinoic acid. *Dev. Biol.* **170,** 243–247 (1995).
5. Hoffman, L., Miles, J., Avaron, F., Laforest, L. & Akimenko, M. A. Exogenous retinoic acid induces a stage-specific, transient and progressive extension of *Sonic hedgehog* expression across the pectoral fin bud of zebrafish. *Int. J. Dev. Biol.* **46,** 949–956 (2002).
6. Tickle, C., Alberts, B., Wolpert, L. & Lee, J. Local application of retinoic acid to the limb bud mimics the action of the polarizing region. *Nature* **296,** 564–566 (1982).
7. Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M. & Shiroishi, T. Elimination of a long-range *cis*-regulatory module causes complete loss of limb-specific *Shh* expression and truncation of the mouse limb. *Development* **132,** 797–803 (2005).
8. Lettice, L. A. *et al.* A long-range *Shh* enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12,** 1725–1735 (2003).
9. Sagai, T. *et al.* Phylogenetic conservation of a limb-specific, *cis*-acting regulator of *Sonic hedgehog* (*Shh*). *Mamm. Genome* **15,** 23–34 (2004).
10. Maas, S. A. & Fallon, J. F. Single base pair change in the long-range *Sonic hedgehog* limb specific enhancer is a genetic basis for preaxial polydactyly. *Dev. Dyn.* **232,** 345–348 (2005).
11. Ballard, W. W., Mellinger, J. & Lechenault, H. A series of normal stages for development of *Scyliorhinus canicula*, the lesser spotted dogfish (*Chondrichthyes*: *Scyliorhinidae*). *J. Exp. Zool.* **267,** 318–336 (1993).
12. Pearse, R. V., Vogan, K. J. & Tabin, C. J. *Ptc1* and *Ptc2* transcripts provide distinct readouts of Hedgehog signaling activity during chick embryogenesis. *Dev. Biol.* **239,** 15–29 (2001).
13. Vandersea, M. W., Fleming, P., McCarthy, R. A. & Smith, D. G. Fin duplications and deletions induced by disruption of retinoic acid signaling. *Dev. Genes Evol.* **208,** 61–68 (1998).
14. Chiang, C. *et al.* Manifestation of the limb prepattern: limb development in the absence of *sonic hedgehog* function. *Dev. Biol.* **236,** 421–435 (2001).
15. Ros, M. A. *et al.* The chick *oligozeugodactyly* (*ozd*) mutant lacks *sonic hedgehog* function in the limb. *Development* **130,** 527–537 (2003).
16. Gritli-Linde, A., Lewis, P., McMahon, A. P. & Linde, A. The whereabouts of a morphogen: direct evidence for short- and graded long-range activity of hedgehog signaling peptides. *Dev. Biol.* **236,** 364–386 (2001).
17. Saunders, J. W. & Gasseling, M. T. *Epithelial Mesenchymal Interactions* (eds Fleischmajer, A. and Billingham, R. E.) 78–97 (William and Wilkins, Baltimore, 1968).
18. Davis, M. C., Shubin, N. H. & Force, A. Pectoral fin and girdle development in the basal actinopterygians *Polyodon spathula* and *Acipenser transmontanus*. *J. Morphol.* **262,** 608–628 (2004).
19. Bemis, W. E. & Grande, L. Early development of the actinopterygian head. I. External development and staging of the paddlefish *Polyodon spathula*. *J. Morphol.* **213,** 47–83 (1992).
20. Prince, V. E., Joly, L., Ekker, M. & Ho, R. K. Zebrafish hox genes: genomic organization and modified colinear expression patterns in the trunk. *Development* **125,** 407–420 (1998).
21. Dahn, R. D. & Fallon, J. F. Interdigital regulation of digit identity and homeotic transformation by modulated BMP signaling. *Science* **289,** 438–441 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Tullis of The University of Puget Sound, Marine Biological Laboratories at Woods Hole, K. Riley of Chicago Aquarium and Pond for specimens, K. Tamura and T. Sagai for sharing unpublished data, P. Beachy for assistance, and K. Monoyios for illustration and graphics assistance. This work was supported by fellowships from The National Institutes of Health (R.D.D.), The Jane Coffin Childs Memorial Fund for Medical Research (W.N.P.), and the Division of Biological Sciences at The University of Chicago (N.H.S.).

Author Contributions All experiments were designed and data generated by R.D.D., except for isolation of full-length skate *Shh* cDNA (W.N.P.) and the *Polyodon* data (M.C.D.). Data were analysed and the manuscript prepared by R.D.D. and N.H.S.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to N.H.S. (nshubin@uchicago.edu).

©2007 Nature Publishing Group

LETTERS

Categorization of behavioural sequences in the prefrontal cortex

Keisetsu Shima[1], Masaki Isoda[1], Hajime Mushiake[1] & Jun Tanji[1,2]

Although it has long been thought that the prefrontal cortex of primates is involved in the integrative regulation of behaviours[1–4], the neural architecture underlying specific aspects of cognitive behavioural planning has yet to be clarified[5–8]. If subjects are required to remember a large number of complex motor sequences and plan to execute each of them individually, categorization of the sequences according to the specific temporal structure inherent in each subset of sequences serves to facilitate higher-order planning based on memory. Here we show, using these requirements, that cells in the lateral prefrontal cortex selectively exhibit activity for a specific category of behavioural sequences, and that categories of behaviours, embodied by different types of movement sequences, are represented in prefrontal cells during the process of planning. This cellular activity implies the generation of neural representations capable of storing structured event complexes at an abstract level, exemplifying the development of macro-structured action knowledge in the lateral prefrontal cortex[9].

We trained two monkeys to perform a series of four movements in 11 different temporal orders[10] (Supplementary Methods). The temporal structures of the 11 motor sequences were designed as follows: four were 'paired' sequences, which included two pairs of movements (for example, turn–turn–push–push), four were 'alternate' sequences, which were composed of an alternation of two movements (for example, pull–turn–pull–turn), and the remaining three were 'four-repeat' sequences, which required the same movement to be repeated four times. After a series of learning trials under visual-signal guidance, the monkeys prepared to perform a sequence of four movements based on memorized information about temporal sequences that fell into the three categories. We recorded from the lateral prefrontal cortex above and below the principal sulcus in the left hemisphere[11]. Among the 376 prefrontal (PF) cells that were task-related, 165 cells exhibited modulation of their activity during the waiting period preceding the first movement-trigger signal; these were the focus of the present study.

Our most striking finding was that in more than half of the cells ($n = 85$) that showed modulated activity during the waiting period, the activity was selectively induced as the monkey prepared to perform sequences from one of the three sequence categories. Figure 1 shows three examples of PF cells that showed category-selective activity. The cell shown in Fig. 1a was preferentially active as the monkey prepared to perform the sequences of motor activities categorized as paired. The activity of this neuron was not significantly different during the periods preceding each of the four paired sequences. In contrast, the activity of the cell shown in Fig. 1b significantly increased when the forthcoming sequences of motor activities were categorized as alternate. The activity of the cell shown in Fig. 1c was selectively elevated when the sequence belonged to the four-repeat category. Using linear regression analysis, we found that 31 cells were selective for the alternate sequences, 32 cells were selective for the paired sequences, and the remaining 22 cells were selective for the four-repeat sequences. In contrast, a minority of the PF cells examined ($n = 30$) was found to be selectively active before the execution of an individual sequence of motor activities (that is, selectivity for any one of the 11 sequences). These 30 cells were not classified as category selective. The remaining 50 cells were not selective either for the category or the sequence.

To what extent do the 85 category-selective cells differentiate between the different categories of sequence during the preparatory period? To answer this question, we plotted a time-varying discharge density function for the populations of cells that responded in preparation for each of the three different categories of sequences. As can be seen in the population data presented in Fig. 2a, activity in cells selective for each of the three sequence categories steadily increased during the preparatory period, clearly showing whether or not the forthcoming sequence belonged to the preferred category of sequences for those cells. Further analysis of the activity of individual cells revealed that differences in the preparatory activity for sequences belonging to one category (within-category variations) were small, whereas across-category variations were large (Supplementary Figs 3 and 4).

Although the monkeys did not make many mistakes, there was unusual cellular activity during error trials. Figure 2b illustrates a typical example of activity that accompanied errors. As long as the performance was correct, the activity of the PF cell shown here increased selectively when the monkey was preparing to perform alternate sequences but not others (black traces). The increase in activity was not observed when the monkey made the mistake of preparing a sequence that belonged to the paired or four-repeat categories (trials labelled 'across category'). In contrast, even when the monkey was preparing to perform an incorrect sequence the activity of the cell was greater if the forthcoming sequence belonged to the alternate category (trials labelled 'within category'). In addition, Fig. 2b (middle panel) illustrates that the activity of the same cell increased when the monkey made the mistake of preparing to perform an alternate sequence when a paired sequence was required. This example is typical in showing that the activity of a cell during the preparatory period faithfully reflected whether or not a sequence from a particular category was planned (see Supplementary Information for other cells).

Subsequently, we examined the extent to which the activity of PF cells during the preparatory period predicted the sequence category that the animals performed. For this purpose, we determined an index of predictive activity (the receiver operating characteristic index)[12,13] using the cellular activity measured during the early, middle and late portions of the preparatory period. This index estimates

[1]Department of Physiology, Tohoku University School of Medicine, Sendai, 980-8575, Japan. [2]Brain Science Research Center, Tamagawa University Research Institute, Tamagawa-Gakuen, Machida, Tokyo, 194-8610, Japan.

©2007 Nature Publishing Group

the probability that an observer can predict the category that will be performed by the monkey on the basis of the spike rate. As presented in Fig. 3, the activity of PF cells during the early part of the delay period was not a good indicator of the sequence category. However, during the last 0.5 s of the preparatory period the activity of PF cells classified on the basis of their selectivity for one of the three sequence categories served as indicators of whether the monkey would perform a paired (green triangles) or four-repeat (red circles) sequence, or an alternate sequence (orange diamonds). The category selectivity started to grow already during early visually guided trials, and developed steadily during memorized trials (Supplementary Fig. 5).

Histological examination revealed that our recording sites were located in the lateral prefrontal cortex in front of the frontal eye field. Category-selective cells were identified more frequently at penetrations that were dorsal to the principal sulcus than at ventrally located penetrations (Fig. 1d; $P < 0.01$ with the χ^2 test), although no spatially selective distribution of the cells classified as belonging to any of the three categories was found (see Supplementary Information).

Previous reports have shown that cells in the medial motor areas of monkeys are selectively activated on the basis of the sequence of multiple movements[14–19]. Neural coding of the sequence of movements has also been reported in the prefrontal cortex[20,21], suggesting prefrontal involvement in representing subjective knowledge of the action sequence[22]. In line with these reports, we found cellular activity that was selective for the sequence of movements in the present study. However, the cells that were selective for a sequence during the preparatory period were not abundant. The novel finding in this study is the identification of PF cells that are selective for the category of the motor sequence to be performed, rather than the individual sequence itself. The information about the category was not obtained from signals in the external world but was generated internally as a neural representation. In particular, knowledge of a category represented an assembly of information about a set of sequences. In that sense, the representation of the category is of a higher order than the representation of the sequences. We therefore propose that the development of category information represented in PF cells implies the generation of a unit of knowledge that specifies the macrostructure of an event series at an advanced level of unification. This view is consistent with a functional magnetic resonance imaging study[23] reporting the involvement of the lateral PF in action selection on the basis of

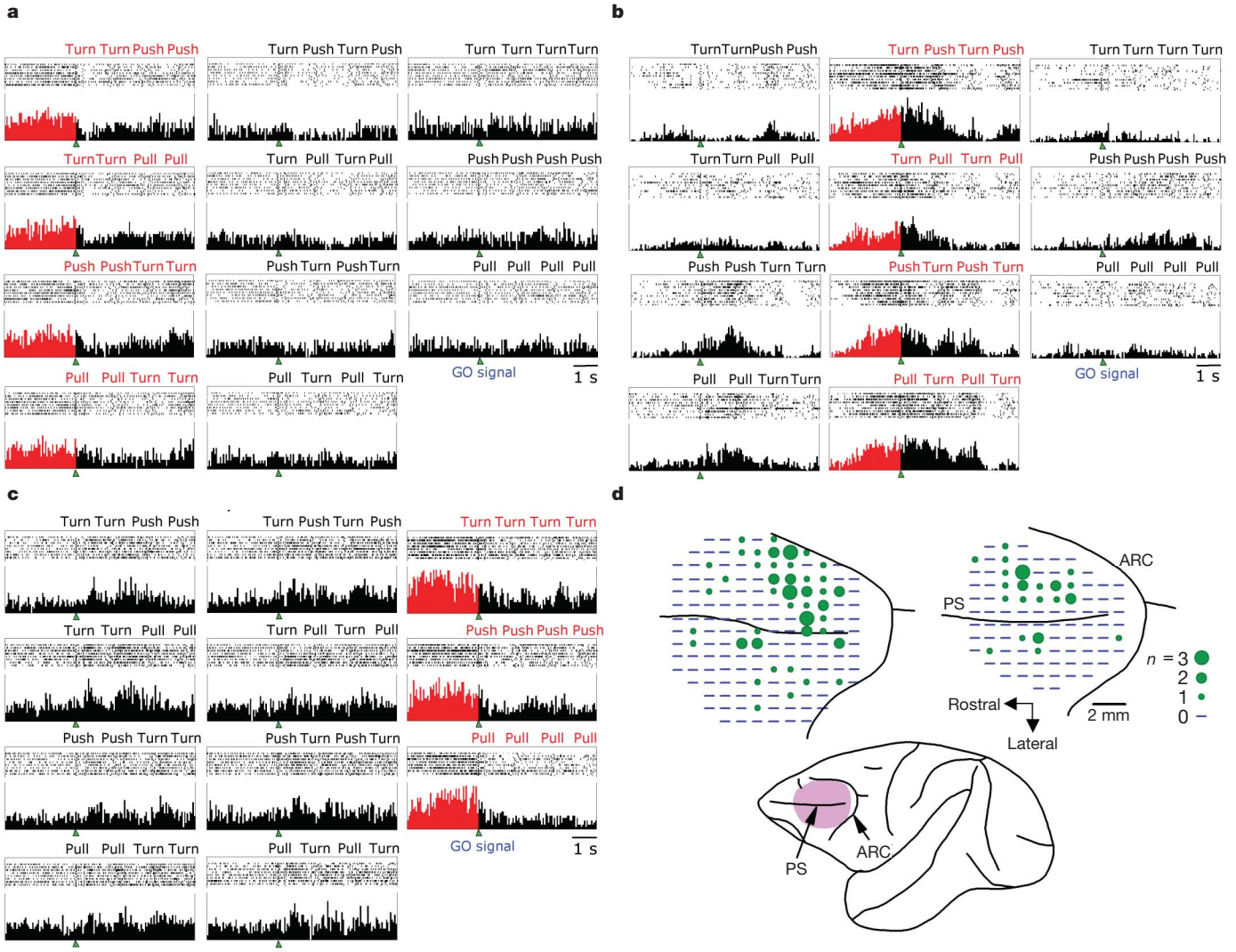


Figure 1 | Activity of PF cells selective for a category of sequences during planning. **a**, Raster displays and peri-event histograms illustrating the cellular activity selective for the 'paired' category. The displays are aligned on the appearance of the GO signal (green-filled triangle) for the first of the memorized movements. **b**, Activity selective for the 'alternate' category. **c**, Activity selective for the 'four-repeat' category. **d**, Top: recording sites of category-selective cells. The size of the circle is proportional to the number of selective cells at each site. Bottom: a cortical surface map showing the surveyed area. PS, principal sulcus; ARC, arcuate sulcus.

©2007 Nature Publishing Group

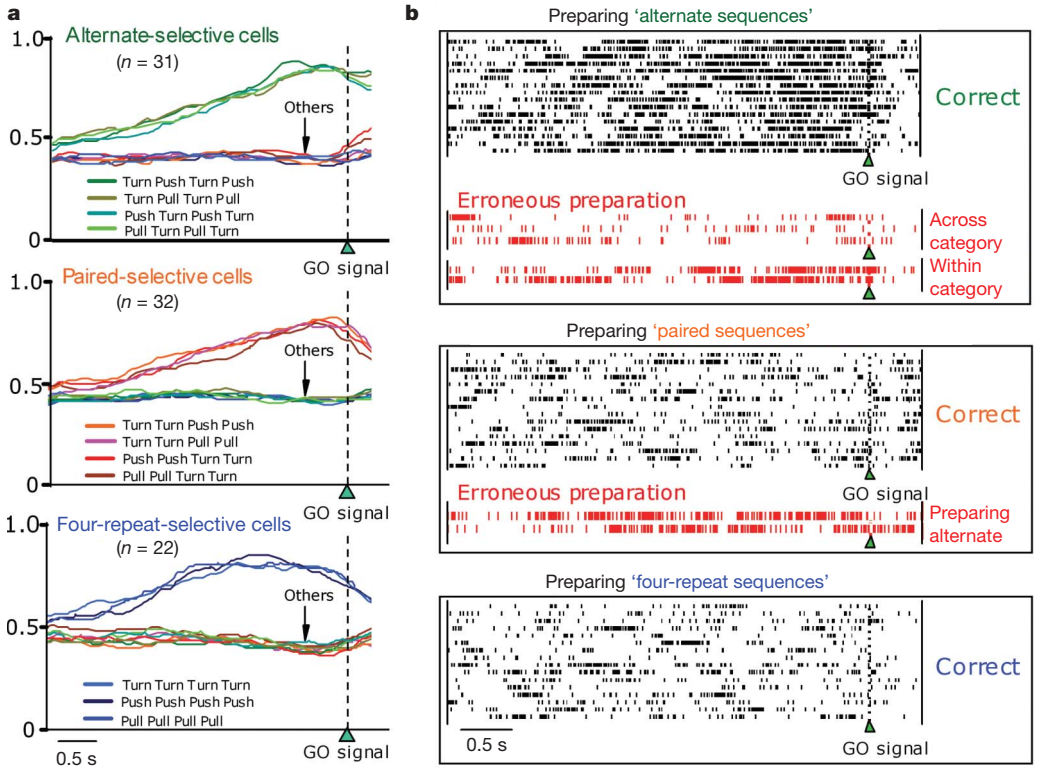


Figure 2 | Time courses of the activity of cell populations showing category selectivity and a comparison of cellular activity before the monkey executed the correct or an incorrect sequence. **a**, Spike density function for populations of PF cells that were selectively active as the monkey prepared behavioural sequences from one of the three categories. The traces compare the activity during preparation for alternate (top), paired (middle) and four-repeat (bottom) sequences with the activity during preparation for sequences of other categories. **b**, Activity of a PF cell when the monkey was preparing to perform the correct or an incorrect behavioural sequence.

hierarchical structures of behavioural plans, and is also consistent with clinical reports that patients with prefrontal lesions are impaired in analysing clusters of action sequences[24] or in formulating action plans that accord with managerial knowledge[25,26]. It is remarkable that about half of the cells modulated during the preparatory period were tuned to a category of sequences. This result is in line with an adaptive coding hypothesis[27] that PF cells adapt their properties to carry information that is relevant to current concerns. It is conjectured that the monkeys used the category information for task performance because they made more within-category than between-category errors, and also because cellular activity during the preparatory period predicted the types of errors that the monkeys would make. This implies that our monkeys were likely to have learned abstract as well as surface structure of movement sequences[28] during extensive training. The responses of PF cells in monkeys trained to categorize animal forms as either doglike or catlike were examined with the use of the technique of 'morphing' the basic form of one animal into the other[29]. The authors found that PF cells selectively responded to visual stimuli that belonged to either the cat or the dog category, showing that category information in the perceptual domain is represented in the prefrontal cortex[30]. Our present results extend these findings on the representation of categories into the domain of behavioural guidance with respect to macro-level planning. Taken together, category-selective cellular activity is useful for

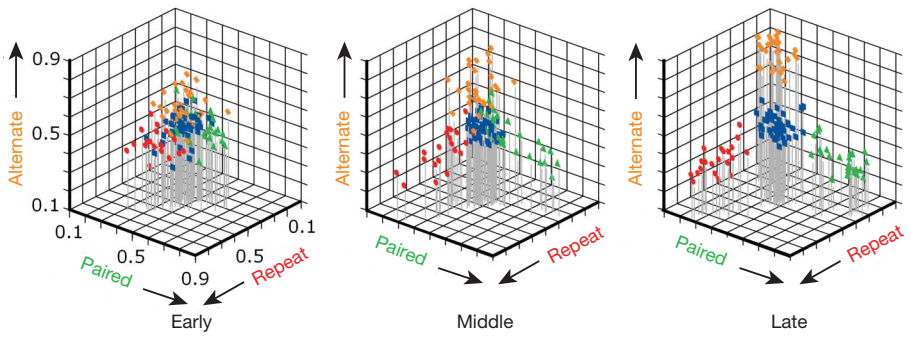


Figure 3 | Plots of the predictive activity of 135 PF cells. The predictive index approximated the accuracy with which an observer could estimate the category of the sequence that the monkey was preparing to execute on the basis of cellular activity during the early (first 500 ms, left), middle (1,000–1,500 ms, middle) and late (last 500 ms, right) phases of the preparatory period. The indices plotted along the *x*, *y* and *z* axes predict the preparation to perform sequences belonging to the three categories shown. Orange diamonds, data for alternate-selective cells; green triangles, paired-selective cells; red circles, four-repeat-selective cells; blue rectangles, non-selective cells.

©2007 Nature Publishing Group

generalizing perceived or internally generated information, for the purpose of guiding behaviour by integrated knowledge represented in the prefrontal cortex.

METHODS

Behavioural procedures. We trained two monkeys (*Macaca fuscata*) to perform a series of four movements in 11 different temporal orders (see Supplementary Fig. 2). Each movement was a push, a pull, or a turn of a handle with the right hand. The task began when the monkey held the handle in a neutral position and waited until a tone signal triggered the first of the four movements. After completion of an individual movement the animal had to wait for the next movement-trigger signal. Initially, during the learning phase, green, yellow and red light-emitting diodes indicated that the monkey was required to push, pull and turn the handle, respectively. The animals underwent five learning trials and subsequently performed the sequential motor task without the visual signals. After five trials based on memory, the sequence of the movements was changed for the next set of trials. For the temporal structures of the 11 motor sequences, the four paired sequences consisted of two movement pairs, four alternate sequences consisted of the alternation of two movements, and the three remaining four-repeat sequences required the one movement to be repeated four times. Thus, after the learning trials, monkeys had to prepare to perform a sequence of four movements on the basis of memorized information about the temporal patterns that fell into one of three categories. When five trials were accomplished under memory, a new sequence began for the next visually guided trials. We used conventional electrophysiological techniques to obtain *in vivo* single-cell recordings[10] from the lateral prefrontal cortex above and below the principal sulcus in the left hemisphere, rostral to the frontal eye field.

Data analysis. We focused on studying cells that had their activity modulated during the preparatory period. To assess statistically how the three categories of motor sequence were related to the activity of the cells, we performed linear regression analysis with the use of the following regression model: firing rate = $\beta_0 + \beta_1 \times$ (categories), where β_0 is the intercept and β_1 the regression coefficient. The categorical factor was the sequence categories (paired, alternate or four-repeat). Subsequently, for cells satisfying this criterion, we examined whether the cellular activity was related to an individual sequence of movements. For this purpose we performed a second round of linear regression analysis with the following equation: firing rate = $\beta_0 + \beta_2 \times$ (sequences), where β_2 is the regression coefficient, and the categorical factors for sequences were the 11 different movement sequences. In this study we defined the activity of a cell to be selectively reflecting a category if it was found to be related to the category but not to an individual sequence within this category. Furthermore, we determined an index that described the association between cellular activity during the preparatory period and the monkey's subsequent selection of a sequence from one of the three sequence categories. This index approximated the ability of an observer to predict the monkey's behaviour from the cellular activity. We performed a receiver operating characteristic analysis[12,13] to obtain this predictive index.

Received 6 July; accepted 21 November 2006.
Published online 20 December 2006.

1. Fuster, J. M. *The Prefrontal Cortex* (Raven, New York, 1989).
2. Goldman-Rakic, P. S. in *Handbook of Physiology: The Nervous System* (ed. Kandel, E. R.) 373–417 (American Physiological Society, Bethesda, Maryland, 1987).
3. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal function. *Annu. Rev. Neurosci.* **24,** 167–202 (2001).
4. Shallice, T. & Burgess, P. W. The domain of supervisory processes and temporal organization of behaviour. *Phil. Trans. R. Soc. Lond. B* **351,** 1405–1411 (1996).
5. Grafman, J. Similarities and distinctions among current models of prefrontal cortical functions. *Ann. NY Acad. Sci.* **769,** 337–368 (1995).
6. Passingham, P. E. *The Frontal Lobes and Voluntary Action* (Oxford Univ. Press, Oxford, 1993).
7. Tanji, J. & Hoshi, E. Behavioral planning in the prefrontal cortex. *Curr. Opin. Neurobiol.* **11,** 164–170 (2001).
8. Koechlin, E., Ody, C. & Kouneiher, F. The architecture of cognitive control in the human prefrontal cortex. *Science* **302,** 1181–1185 (2003).
9. Wood, J. N. & Grafman, J. Human prefrontal cortex: processing and representational perspectives. *Nature Rev. Neurosci.* **4,** 139–147 (2003).
10. Shima, K. & Tanji, J. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *J. Neurophysiol.* **84,** 2148–2160 (2000).
11. Saito, N., Mushiake, H. & Tanji, J. Representation of immediate and final behavioral goals in the monkey prefrontal cortex during an instructed delay period. *Cereb. Cortex* **15,** 1535–1546 (2005).
12. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
13. Nieder, A. & Miller, E. K. A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl Acad. Sci. USA* **101,** 7457–7462 (2004).
14. Mushiake, H., Inase, M. & Tanji, J. Selective coding of motor sequence in the supplementary motor area of the monkey cerebral cortex. *Exp. Brain Res.* **82,** 208–210 (1990).
15. Nakamura, K., Sakai, K. & Hikosaka, O. Neuronal activity in medial frontal cortex during learning of sequential procedures. *J. Neurophysiol.* **80,** 2671–2687 (1998).
16. Tanji, J. & Shima, K. Role for supplementary motor area cells in planning several movements ahead. *Nature* **371,** 413–416 (1994).
17. Tanji, J. Sequential organization of multiple movements: involvement of cortical motor areas. *Annu. Rev. Neurosci.* **24,** 631–651 (2001).
18. Procyk, E., Tanaka, Y. L. & Joseph, J. P. Anterior cingulate activity during routine and nonroutine sequential behaviors in macaques. *Nature Neurosci.* **3,** 502–508 (2000).
19. Isoda, M. & Tanji, J. Contrasting neuronal activity in the supplementary and frontal eye fields during temporal organization of multiple saccades. *J. Neurophysiol.* **90,** 3054–3065 (2003).
20. Barone, P. & Joseph, J. P. Prefrontal cortex and spatial sequencing in macaque monkey. *Exp. Brain Res.* **78,** 447–464 (1989).
21. Averbeck, B. B., Chafee, M. V., Crowe, D. A. & Georgopoulos, A. P. Neural activity in prefrontal cortex during copying geometrical shapes. I. Single cells encode shape, sequence, and metric parameters. *Exp. Brain Res.* **150,** 127–141 (2003).
22. Averbeck, B. B., Sohn, J. W. & Lee, D. Activity in prefrontal cortex during dynamic selection of action sequences. *Nature Neurosci.* **9,** 276–282 (2006).
23. Koechlin, E. & Jubault, T. Broca's area and the hierarchical organization of human behavior. *Neuron* **50,** 963–974 (2006).
24. Zalla, T., Pradat-Diehl, P. & Sirigu, A. Perception of action boundaries in patients with frontal lobe damage. *Neuropsychologia* **41,** 1619–1627 (2003).
25. Sirigu, A. *et al.* Planning and script analysis following prefrontal lobe lesions. *Ann. NY Acad. Sci.* **769,** 277–288 (1995).
26. Sirigu, A. *et al.* Selective impairments in managerial knowledge following pre-frontal cortex damage. *Cortex* **31,** 301–316 (1995).
27. Duncan, J. An adaptive coding model of neural function in prefrontal cortex. *Nature Rev. Neurosci.* **2,** 820–829 (2001).
28. Dominey, P. F., Lelekov, T., Ventre-Dominey, J. & Jeannerod, M. Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences. *J. Cogn. Neurosci.* **10,** 734–751 (1998).
29. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller, E. K. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291,** 312–316 (2001).
30. Miller, E. K., Freedman, D. J. & Wallis, J. D. The prefrontal cortex: categories, concepts and cognition. *Phil. Trans. R. Soc. Lond. B* **357,** 1123–1136 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Author Contributions K.S. and J.T. performed all aspects of the study, including the design of the experiment, collecting and analysing the data, and writing the manuscript. M.I. assisted in experimental design and in collecting the data. H.M. assisted in data analysis and manuscript preparation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.T. (tanji@lab.tamagawa.ac.jp).

©2007 Nature Publishing Group

LETTERS

Aberrant innate immune response in lethal infection of macaques with the 1918 influenza virus

Darwyn Kobasa[1], Steven M. Jones[2,3], Kyoko Shinya[5], John C. Kash[6], John Copps[8], Hideki Ebihara[2,9,10,11], Yasuko Hatta[12], Jin Hyun Kim[12], Peter Halfmann[12], Masato Hatta[12], Friederike Feldmann[2], Judie B. Alimonti[2], Lisa Fernando[2], Yan Li[1], Michael G. Katze[6,7], Heinz Feldmann[2,4] & Yoshihiro Kawaoka[9,10,11,12]

The 1918 influenza pandemic was unusually severe, resulting in about 50 million deaths worldwide[1]. The 1918 virus is also highly pathogenic in mice, and studies have identified a multigenic origin of this virulent phenotype in mice[2–4]. However, these initial characterizations of the 1918 virus did not address the question of its pathogenic potential in primates. Here we demonstrate that the 1918 virus caused a highly pathogenic respiratory infection in a cynomolgus macaque model that culminated in acute respiratory distress and a fatal outcome. Furthermore, infected animals mounted an immune response, characterized by dysregulation of the antiviral response, that was insufficient for protection, indicating that atypical host innate immune responses may contribute to lethality. The ability of influenza viruses to modulate host immune responses, such as that demonstrated for the avian H5N1 influenza viruses[5], may be a feature shared by the virulent influenza viruses.

Genes of the 1918 virus were constructed from published sequences[6–11], and the virus was generated using plasmid-based reverse genetics, as previously described[12]. The virus was highly virulent in intranasally inoculated mice with an LD_{50} (dose required to kill 50% of animals) of $10^{3.5}$ plaque-forming units (PFU), comparable to that previously reported[4]. Cynomolgus macaques (*Macaca fascicularis*), previously used to study H5N1 influenza virus pathogenesis[13,14], were selected as a nonhuman primate model for further analysis of virulence. Macaques ($n = 7$) were infected with the 1918 virus) or a conventional human virus, A/Kawasaki/173/01 (K173; H1N1) ($n = 3$), via multiple routes, as in the H5N1 virus experiment[13]. All 1918-virus-infected animals became symptomatic within 24 h post-infection. They appeared depressed, were hesitant to eat or drink normal food items, and showed respiratory complications such as nasal discharge and non-productive cough. They became progressively more debilitated and eventually developed an acute respiratory distress syndrome. Two macaques infected with the 1918 virus and one with K173 were euthanized on each of days 3 and 6 for analysis. Of these, one 1918-virus-infected animal had reached the predetermined score for euthanasia on day 6. The remaining animals, originally scheduled for euthanasia on day 21 post-infection, were euthanized on day 8 owing to severity of symptoms in 1918-virus-infected animals.

In 1918-virus-infected animals respiratory signs were the most pronounced indication of illness with an increase in respiration rate, from a post-infection average of 26.4, to a range of 72–84 on day 8. Decreases in lung function shown by a decrease in blood oxygen saturation of as much as 36% compared with pre-infection levels were detected by pulse oximetry. Consistent changes in heart rate and blood pressure were not observed. In contrast, animals infected with the K173 virus showed few, very mild, clinical signs.

The 1918 virus was present at high titres in both the upper and lower respiratory tissues on days 3, 6 and 8, whereas the K173 virus was isolated mainly from the upper respiratory tissues on days 3 and 6 at appreciably lower titres and on day 8 in only one tonsil (Fig. 1). The 1918 virus was recovered from the heart and spleen of some animals, but neither virus was isolated from the brain, kidneys, liver or colon of any animal. Virus was detected more efficiently with nasal or throat swabs from 1918-virus-infected animals than for K173-virus-infected animals (Supplementary Table 1), but neither virus was isolated in rectal or genital swabs, or in the blood, on any day.

Lungs of the 1918-virus-infected macaques were the only tissue to exhibit macroscopic pathologic changes as severe lesions, with 60% to 90% of the lung tissue affected by days 6–8 (Supplementary Fig. 1). Profuse watery and bloody liquid filled infected areas, greatly reducing lung function. On day 3 with either virus, there was evidence of some degree of alveolar damage with concomitant detection of viral antigen (Supplementary Fig. 2a–i); however, K173 was not isolated from the lungs, indicating low levels of replication. In contrast to lesions of the K173-infected animal, lungs of 1918-virus-infected animals showed infection of multifarious cells of the alveolar wall on day 3 (Supplementary Fig. 2f, g). Common to infection with either the K173 or 1918 virus were flattened linear alveolar cells that were positive for viral antigen without showing desquamation into the alveolar space (Supplementary Fig. 2d, f). A prominent characteristic of 1918-virus-infected lungs was antigen in plump alveolar cells and desquamation of these cells into the alveolar space (Supplementary Fig. 2g, h). These differences in the early phase of infection between the 1918 and K173 viruses might account for later differences in inflammation and the eventual pathologic state of the lungs.

By days 6 and 8, the lungs of K173-virus-infected animals showed signs of healing, evidenced by thickening of the alveolar wall and no viral antigen expression (Fig. 2a–c, day 8; day 6 not shown). In contrast, the lungs of all 1918-virus-infected animals showed worsening alveolar damage and substantial viral antigen (Fig. 2d–h). Extensive oedema and haemorrhagic exudates were prominent (Fig. 2d–f), as reported for patients who succumbed to the 'Spanish' influenza[15]. By day 8, expression of viral antigen was diminished

[1]Respiratory Viruses, and [2]Special Pathogens Program National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba R3E 3R2, Canada. [3]Department of Immunology and [4]Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba R3E 3R2, Canada. [5]The Avian Zoonosis Research Centre, Tottori University, Tottori 680-8550, Japan. [6]Department of Microbiology, School of Medicine, and [7]Washington National Primate Research Center, University of Washington, Seattle, Washington 98195, USA. [8]National Centre for Foreign Animal Diseases, Canadian Food Inspection Agency, Canadian Science Centre for Human and Animal Health, Winnipeg, Manitoba R3E 3M4, Canada. [9]Division of Virology, Department of Microbiology and Immunology and [10]International Research Center for Infectious Diseases, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan. [11]CREST, Japan Science and Technology Agency, Saitama 322-0012, Japan. [12]Department of Pathobiological Sciences, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA.

©2007 Nature Publishing Group

or undetectable in consolidated alveolar areas, but still appreciable around consolidated areas (Fig. 2g). Bronchiolitis and bronchitis with expression of viral antigen were prominent at this time (Fig. 2h).

Among the chemokines and cytokines tested (see Methods), substantial increases of interleukin-6 (IL-6), IL-8, CCL2 (monocyte chemotactic protein-1) and CCL5 (RANTES) were detected in the sera of infected animals compared with pre-infection levels. No changes in IL-2, IL-4, IFNγ or TNF-α were detected. Both IL-8 (day 6) and CCL5 (days 6 and 8) were elevated to similar levels in animals infected with either virus (Fig. 3). CCL2 was elevated 1.4- to 1.8-fold in K173-infected animals on days 3 and 6, respectively, whereas in the 1918-virus-infected animals, increases were higher (2- to 4-fold), but did not achieve statistical significance. The most striking changes were seen in IL-6. On day 0, IL-6 in all macaques ranged from 2.85 to 13.04 pg ml^{-1}, except for one animal with a baseline measurement of 251.32 pg ml^{-1}. In 1918-virus-infected animals, IL-6 increased 3- to 9-fold by day 3 post-infection ($P = 0.10$), 6- to 19-fold by day 6 ($P = 0.07$), and 5- to 25-fold by day 8, while remaining throughout the study between 2.85 and 4.59 pg ml^{-1} in the K173-infected animals. In humans experimentally infected with influenza virus, IL-6

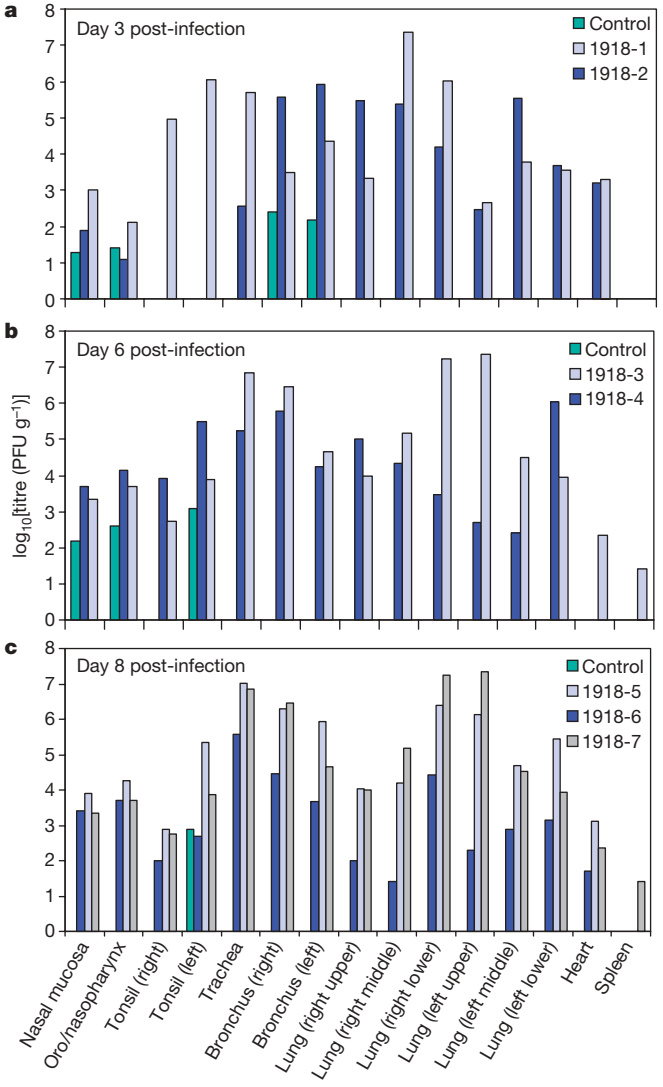


Figure 1 | Viral replication in nonhuman primate tissues. **a**–**c**, Virus titres were determined in plaque assays of tissue homogenates on days 3 (**a**), 6 (**b**) and 8 (**c**) post-infection in animals infected with the K173 (control) or 1918 virus (1918-1 to 1918-7). Individual titres for each tissue in each animal are reported as $\log_{10}$ (PFU per g tissue). Virus was not isolated from the colon, liver, kidneys or brain of any animal.

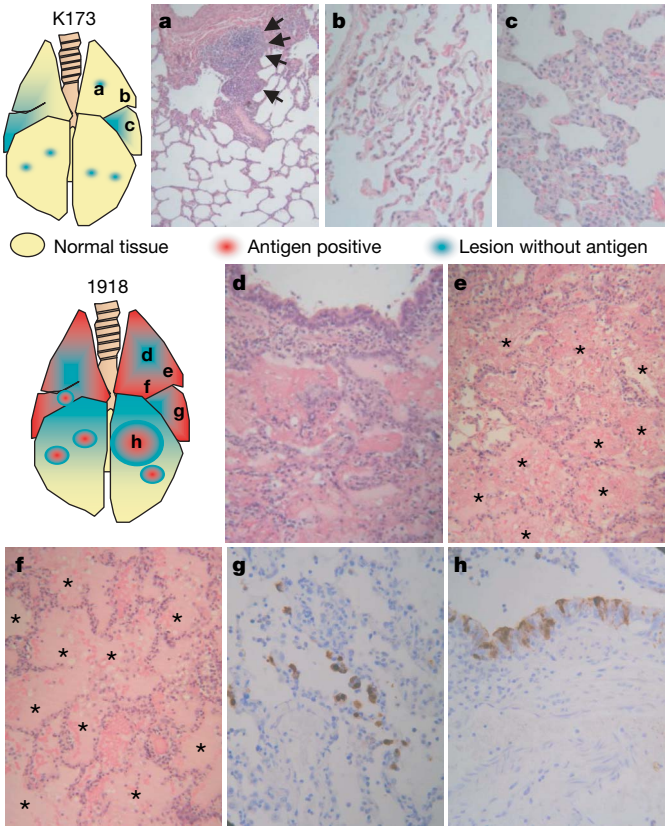


Figure 2 | Pathologic examination of lungs infected with the 1918 or K173 virus on day 8 post-infection. **a**, In lungs of a K173-virus-infected animal, peribronchiolitis with lymph follicle formation was detected throughout the lung (arrows). **b**, **c**, Mild thickening of the alveolar wall was observed in the middle lobes, but antigen was not detected. **d**–**f**, Most areas of the lungs of a 1918-virus-infected animal contained consolidated lesions (**d**) consisting of bronchiolitis and alveolitis (**e**) with fibrinous and inflammatory exudates (*) and alveolar oedema (**f**) with proteinaceous fluid and haemorrhage (*). **g**, **h**, Viral antigen (brown) was detected in the large regenerative alveolar cells (**g**) and the bronchiolar epithelial cells in lesions (**h**). Magnifications are ×25 in **a**; ×240 in **b**, **c**, **g** and **h**; and ×120 in **d**–**f**.

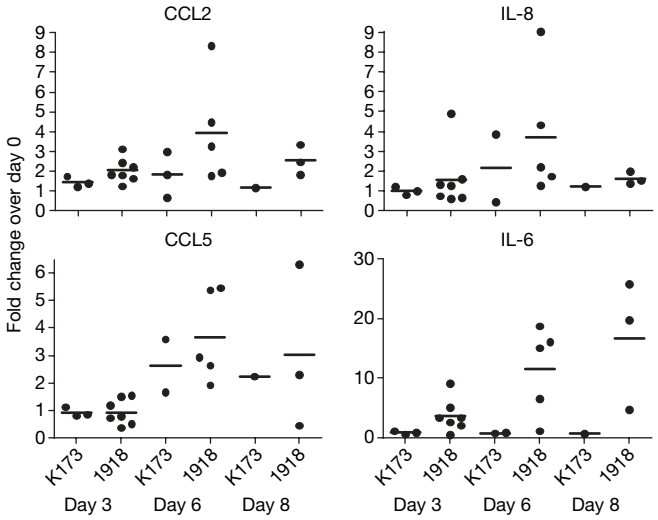
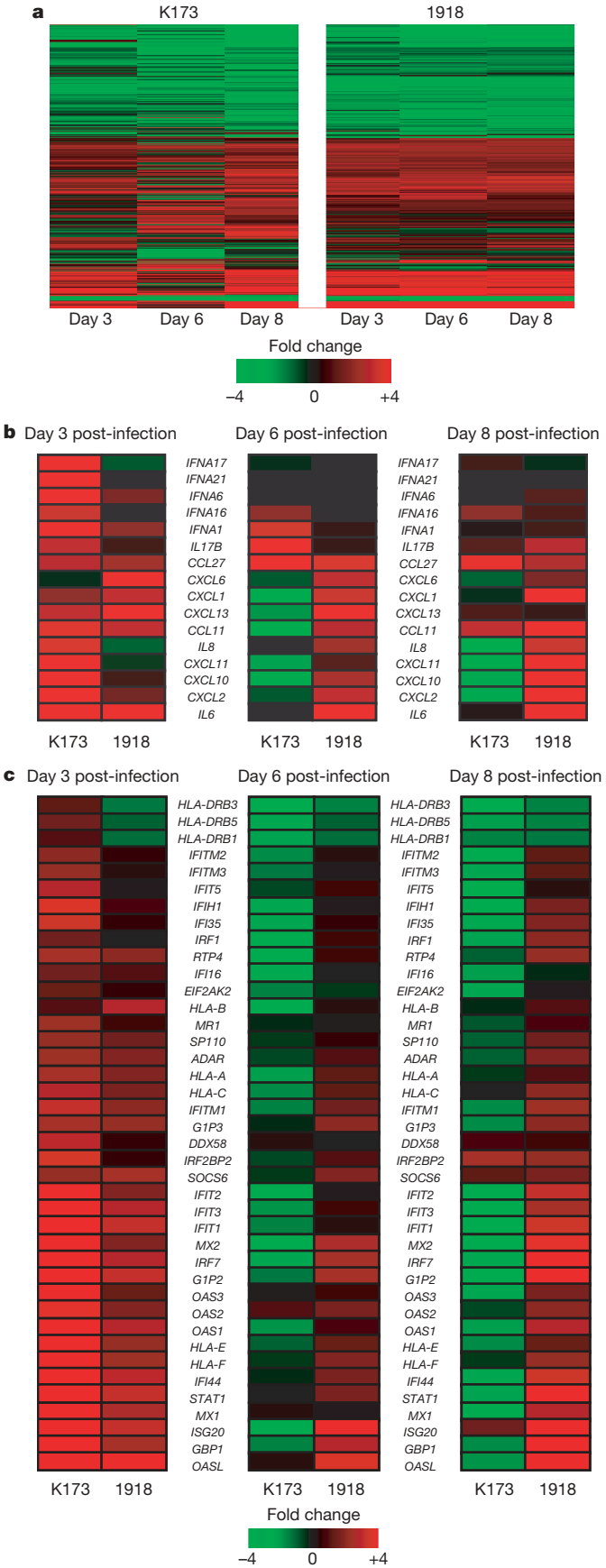


Figure 3 | Chemokine and cytokine levels in serum. The levels of chemokines (CCL2 and CCL5) and cytokines (IL-6 and IL-8) in serum were analysed by a cytometric bead array assay. The levels on days 3, 6 and 8 were compared with the day 0 baseline to determine the relative changes within each animal. The bars represent the average for each group.

©2007 Nature Publishing Group



expression at both the site of replication and in sera directly correlated with the extent of viral shedding and fever and is thought to have a role in mediating the clinical manifestations of infection[16,17]. A similar relationship may also explain the progression to severe clinical signs in macaques infected with the 1918 virus.

To investigate the regulation of the host response to the 1918 virus, we considered the global gene expression profiles in bronchi, a lower respiratory tissue with substantial replication of both viruses (see Fig. 1a–c), using microarray analysis (Fig. 4a, b). In general, overall gene expression patterns of individual 1918-virus-infected animals were similar, particularly on days 6–8 (Supplementary Fig. 3). One striking observation was the relative constancy in gene expression profiles in the 1918-virus-infected animals at days 3 through 8 post-infection, in marked contrast to the K173 virus-infected animals, which showed an appreciably more dynamic response during the course of infection. The sustained host response in 1918-virus-infected macaques is similar to that recently reported in 1918-virus-infected mice[18], indicating that critical decisions influencing the outcome of the infection may occur very early, as recently suggested for H5N1 virus infection in humans[19]. The functional response associated with the observed gene expression changes was analysed by gene ontology (see Methods). This analysis demonstrated that activation of immune-response-related genes in K173-infected animals was highest at day 3 post-infection, whereas the expression of these genes was relatively constant throughout the course of 1918 virus infection (Supplementary Fig. 4). Interestingly, we observed a dramatic activation of genes involved in cellular metabolism in the K173-infected animals on days 6 to 8 post-infection, probably due to cell proliferation and tissue remodelling after clearance of the virus.

Detailed analysis of immune-response-related gene expression changes (Fig. 4b) demonstrated that bronchi from both groups of animals showed activation of several inflammatory chemokine and cytokine genes, including *CCL11* and *IL-6*. *IL-6* messenger RNA was expressed at a high level until day 8 post-infection in the 1918-virus-infected animals, but not the K173-infected animals, consistent with levels detected in serum (Fig. 3). Several key cytokine genes, including *IL-8* and *CXCL11*, showed a delay in activation in 1918-virus-infected animals, although several chemokines important for the activation and recruitment of neutrophils, including *CXCL6* and *CXCL1*, were preferentially upregulated. Strikingly, K173-infected animals showed a marked increase in expression of mRNAs for many type I interferons (IFNs) and a corresponding increase in mRNA expression of type-I-IFN-stimulated genes early in infection (Fig. 4c), coinciding with the greatest load of the virus (Fig. 1). This response was downregulated on days 6 and 8 post-infection, when the K173 virus was not detected. 1918-virus-infection, in contrast, induced much fewer IFN-α genes, suggesting that it caused an altered antiviral response in the bronchus. Accordingly, the 1918-virus-infected animals also showed differential activation of type-I-IFN-stimulated gene expression on days 3 to 8, despite viral titres in bronchi that were 10–5,000-fold higher than in K173-virus-infected

Figure 4 | Microarray analysis of gene expression in the bronchi of macaques on days 3, 6 and 8 post-infection. **a**, Overall gene expression profiles in K173- and 1918-virus-infected animals. **b**, Expression of chemokines and cytokines as determined by expression oligonucleotide microarray analysis of bronchial tissue from macaques infected with the K173 or 1918 influenza virus. **c**, Expression of type I interferon (IFN)-stimulated genes in K173- and 1918-virus-infected macaque bronchi. For each K173 virus infection point, the data presented are gene expression changes calculated from technical replicate arrays ($n = 4$). The data shown for the day 3 and day 6 1918 virus infections are error-weighted averages for two animals ($n = 8$ arrays), whereas the day 8 1918 virus infection data are the error-weighted averages for three animals ($n = 12$ arrays). Genes included in this figure showed ≥2-fold ($P < 0.01$) differences in expression in at least one experiment. Genes shown in red were upregulated and those shown in green were downregulated in infected relative to mock-infected animals. See Supplementary Table 2 for full gene names.

©2007 Nature Publishing Group

animals. Although the mechanisms underlying the differences in regulation of type I IFN responses remain elusive, our data suggest that the 1918 virus induces an antiviral response different from that of K173, possibly including reduced sensitivity to type I IFN responses. Finally, IFN-β1 mRNA was either not induced or was down-regulated in all samples except for a modest induction of IFN-β1 in the K173-infected animal at day 3 post-infection (not shown). Only limited IFN-β1 expression might be expected given that IFN-β1 is a first-wave interferon and would probably be induced earlier than our first sample collection (that is, day 3).

An important pathway in activation of the antiviral response to influenza virus infection occurs through the activities of DDX58 (or retinoic-acid-inducible protein I) and IFIH1 (or melanoma differentiation-associated gene 5)[20,21]. Both genes were induced in the K173-virus-infected, but not the 1918-virus-infected, animals (Fig. 4d). The nonstructural protein 1 (NS1) protein of virulent influenza viruses can modulate the IFN-mediated antiviral response[22–24] and RIG-I is a target of NS1 immunosuppressive activity[20,21,25]. Although an immunomodulatory role for 1918 virus NS1 was not shown in mice[8], possibly owing to its species specificity, NS1 does influence the virulence of the H5N1 virus in a mammalian model[5]. This supports the idea that NS1 may be a subject for further investigation as an effector of the aberrations in the macaque immune response to 1918 virus infection and emphasizes the value of development of the macaque model for studying influenza pathogenesis[26].

Results of the present study, the first in nonhuman primates, indicate that atypical expression of the innate immune response may be a critical determinant of the severity and outcome of infection by the 1918 virus. Better understanding of the virus–host interaction will aid development of interventions that can interfere with the virus' ability to modulate host innate immune responses and thus alter outcome of severe infections due to viruses such as the H5N1 viruses circulating at present.

METHODS

Viruses. Genes of the 1918 (GenBank DQ208309, DQ208310, DQ208311, AF117241, AY744935, AF250356, AY130766, AF233238) and K173 virus were constructed with the 5′ and 3′ noncoding sequences of A/WSN/33 (H1N1) and cloned into plasmid vector pPolI, as previously described[3]. The 1918 and K173 viruses were generated by reverse genetics[12], and titred stocks prepared[3], as previously described. All procedures with the 1918 influenza virus were performed in the biosafety level 4 facility of the National Microbiology Laboratory of the Public Health Agency of Canada.

Determination of the lethal dose for the 1918 virus in mice. Isoflurane-anaesthetized 6-week-old female BALB/c mice were intranasally inoculated with 10-fold serial dilutions (five mice per dilution) of virus in 100 μl of phosphate-buffered saline and monitored daily for disease symptoms and survival. The LD_{50} was calculated using the method of Reed and Muench[27].

Viral pathogenicity in nonhuman primates. Ten cynomolgus macaques (*Macaca fascicularis*), 9–19 years old, weighing 3.7–12 kg, were confirmed seronegative against current H1N1 and H3N2 influenza reference viruses by haemagglutination inhibition assay. Animals were infected with K173 ($n = 3$), as a conventional human virus control, and 1918 ($n = 7$) viruses through a combination of intratracheal (4 ml), intranasal (0.5 ml per nostril), intraocular (0.5 ml per eye) and oral (1 ml) routes with suspension containing 10^6 PFU ml^{-1} (infectious dose by all routes is 7×10^6 PFU). Before infection and on days 3, 6 and 8 post-infection, blood and oral, nasal, throat, genital and rectal swabs were collected from anaesthetized animals and suspended in 1 ml of MEM containing 0.3% bovine serum albumin and antibiotics (MEM/BSA). Animals were monitored daily for clinical signs, using an approved scoring sheet, and on days 3, 6 and 8 post-infection vital signs including pulse rate, blood pressure, temperature, respiration rate and blood O_2 saturation, as measured by pulse oximetry, were recorded.

One K173- and two 1918-virus-infected animals were euthanized on day 3 and on day 6 post-infection; all remaining animals were euthanized on day 8 post-infection for complete necropsy. Tissue samples were placed in RNAlater (Ambion) for subsequent RNA extraction (Qiagen RNAlater kit). The remaining tissue was fixed in 10% phosphate-buffered formalin. Fixed tissues were dehydrated, embedded in paraffin, cut into 5-μm-thick sections and stained with standard haematoxylin and eosin. For viral antigen detection, sections were processed for immunostaining by the two-step dextran polymer method (DAKO), with a rabbit polyclonal antibody to WSN.

All animal experiments were performed under an approved animal-use document and according to the guidelines of the Canadian Council on Animal Care.

Expression microarrays. Equal masses of total RNA isolated from bronchi collected from infected macaques were amplified with a Low RNA Input Linear Amplification Kit (5188-5339; Agilent Technology) according to the manufacturer's instructions and as previously described[28]. Global gene expression in infected bronchi was compared to pooled RNA prepared from equal masses of total RNA from whole lung tissue of three uninfected macaques. Mock-infected intact lung was used as a reference for the microarrays because of a lack of mock-infected bronchus tissue. Probe labelling and microarray slide hybridization were performed as previously described[28] with custom rhesus macaque (*Macaca mulatta*) oligonucleotide microarrays containing 22,000 rhesus probes corresponding to ~18,000 unique rhesus genes (designed in collaboration with Agilent Technologies). Raw microarray image files were processed using Feature Extraction 8.1 software (Agilent Technologies) and entered into a custom-designed relational database (Expression Array Manager) and analysed with Rosetta Resolver System 5.1 (Rosetta Biosoftware, Seattle, Washington) and Spotfire Decision Site for Functional Genomics 8.1 (Spotfire, Somerville, Massachusetts). Primary data are available at <http://expression.microslu.washington.edu> in accordance with proposed MIAME standards[29]. Gene ontology analysis was performed using FatiGO (http://fatigo.bioinfo.cipf.es/) as described in reference 30.

Serum cytokine analysis. Serum was obtained from blood collected in K_2EDTA vacutainer tubes (BD Biosciences) that were centrifuged at 1,540*g* for 10 min and γ-irradiated (5 Mrad) for removal from biosafety level 4 containment. Commercial kits for detection of human cytokines and chemokines (IL-2, IL-4, IL-5, IL-6, IFN-γ, TNF-α, IL-8, IL-10, IL-1β, IL12p70, CCL2, CCL5, M1G and IP-10) were screened for cross-reactive detection of homologous macaque proteins. Of these, IL-2, IL-4, IL-6, IL-8, CCL2, TNF-α, IFN-γ and CCL5 exhibited significant cross-reactivity and were detected in supernatants of lipopolysaccharide or phorbol 12-myristate 13-acetate/ionomycin-stimulated macaque peripheral blood mononuclear cells. Levels of each were assayed using the Human Chemokine CBA Kit I and the human IL-6 Flex set (BD Biosciences) according to the manufacturer's instructions. The samples were acquired with the LSRII (BD Biosciences) flow cytometer. Chemokine and IL-6 samples were analysed with BD Biosciences' CBA and FCAP software, respectively.

Virus titration. Tissue homogenates (10% w/v) were prepared in MEM/BSA. Debris was pelleted by centrifugation (2,000*g*, 5 min) and virus titres determined in 10-fold serial dilutions of supernatant by standard plaque assay on MDCK cells, in duplicate for each dilution. Virus was similarly determined in the blood and swab suspensions.

Received 11 September; accepted 29 November 2006.

1. Johnson, N. P. & Mueller, J. Updating the accounts: global mortality of the 1918–1920 "Spanish" influenza pandemic. *Bull. Hist. Med.* **76,** 105–115 (2002).
2. Tumpey, T. M. *et al.* Existing antivirals are effective against influenza viruses with genes from the 1918 pandemic virus. *Proc. Natl Acad. Sci. USA* **99,** 13849–13854 (2002).
3. Kobasa, D. *et al.* Enhanced virulence of influenza A viruses with the haemagglutinin of the 1918 pandemic virus. *Nature* **431,** 703–707 (2004).
4. Tumpey, T. M. *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310,** 77–80 (2005).
5. Seo, S. H., Hoffmann, E. & Webster, R. G. Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. *Nature Med.* **8,** 950–954 (2002).
6. Reid, A. H., Fanning, T. G., Hultin, J. V. & Taubenberger, J. K. Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. *Proc. Natl Acad. Sci. USA* **96,** 1651–1656 (1999).
7. Reid, A. H., Fanning, T. G., Janczewski, T. A. & Taubenberger, J. K. Characterization of the 1918 "Spanish" influenza virus neuraminidase gene. *Proc. Natl Acad. Sci. USA* **97,** 6785–6790 (2000).
8. Basler, C. F. *et al.* Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. *Proc. Natl Acad. Sci. USA* **98,** 2746–2751 (2001).
9. Reid, A. H., Fanning, T. G., Janczewski, T. A., McCall, S. & Taubenberger, J. K. Characterization of the 1918 "Spanish" influenza virus matrix gene segment. *J. Virol.* **76,** 10717–10723 (2002).
10. Reid, A. H., Fanning, T. G., Janczewski, T. A., Lourens, R. M. & Taubenberger, J. K. Novel origin of the 1918 pandemic influenza virus nucleoprotein gene. *J. Virol.* **78,** 12462–12470 (2004).
11. Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437,** 889–893 (2005).
12. Neumann, G. *et al.* Generation of influenza A viruses entirely from cloned cDNAs. *Proc. Natl Acad. Sci. USA* **96,** 9345–9350 (1999).

©2007 Nature Publishing Group

13. Rimmelzwaan, G. F. *et al.* Pathogenesis of influenza A (H5N1) virus infection in a primate model. *J. Virol.* **75,** 6687–6691 (2001).
14. Kuiken, T., Rimmelzwaan, G. F., Van Amerongen, G. & Osterhaus, A. D. Pathology of human influenza A (H5N1) virus infection in cynomolgus macaques (*Macaca fascicularis*). *Vet. Pathol.* **40,** 304–310 (2003).
15. Winternitz, M. C., Wason, I. M. & McNamara, F. P. *The Pathology of Influenza.* (Yale Univ. Press, New Haven, Connecticut, 1920).
16. Hayden, F. G. *et al.* Local and systemic cytokine responses during experimental human influenza A virus infection. Relation to symptom formation and host defense. *J. Clin. Invest.* **101,** 643–649 (1998).
17. Skoner, D. P., Gentile, D. A., Patel, A. & Doyle, W. J. Evidence for cytokine mediation of disease expression in adults experimentally infected with influenza A virus. *J. Infect. Dis.* **180,** 10–14 (1999).
18. Kash, J. C. *et al.* Genomic analysis of increased host immune and cell death responses induced by 1918 influenza virus. *Nature* **443,** 578–581 (2006).
19. de Jong, M. D. *et al.* Fatal outcome of human influenza A (H5N1) is associated with high viral load and hypercytokinemia. *Nature Med.* **12,** 1203–1207 (2006).
20. Matikainen, S. *et al.* Tumor necrosis factor alpha enhances influenza A virus-induced expression of antiviral cytokines by activating RIG-I gene expression. *J. Virol.* **80,** 3515–3522 (2006).
21. Kato, H. *et al.* Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441,** 101–105 (2006).
22. Garcia-Sastre, A. *et al.* Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. *Virology* **252,** 324–330 (1998).
23. Krug, R. M., Yuan, W., Noah, D. L. & Latham, A. G. Intracellular warfare between human influenza viruses and human cells: the roles of the viral NS1 protein. *Virology* **309,** 181–189 (2003).
24. Li, S., Min, J. Y., Krug, R. M. & Sen, G. C. Binding of the influenza A virus NS1 protein to PKR mediates the inhibition of its activation by either PACT or double-stranded RNA. *Virology* **349,** 13–21 (2006).
25. Pichlmair, A. *et al.* RIG-I-mediated antiviral responses to single-stranded RNA bearing 5′ phosphates. *Science* **314,** 997–1001 (2006).
26. Baas, T. *et al.* Integrated molecular signature of disease: analysis of influenza virus-infected macaques through functional genomics and proteomics. *J. Virol.* **80,** 10813–10828 (2006).
27. Reed, L. J. & Muench, H. A simple method of estimating fifty per cent endpoints. *Am. J. Hyg.* **27,** 493–497 (1938).
28. Kash, J. C. *et al.* Global suppression of the host antiviral response by Ebola- and Marburgviruses: increased antagonism of the type I interferon response is associated with enhanced virulence. *J. Virol.* **80,** 3009–3020 (2006).
29. Brazma, A. *et al.* Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nature Genet.* **29,** 365–371 (2001).
30. Al-Shahrour, F., Minguez, P., Vaquerizas, J. M., Conde, L. & Dopazo, J. BABELOMICS: a suite of web tools for functional annotation and analysis of groups of genes in high-throughput experiments. *Nucleic Acids Res.* **33,** W460–W464 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Dick, J. Gren, A. Grolla and P. Melito for help with animal care, and V. Carter, M. Thomas and S. Proll for microarray technical assistance. We also thank J. Gilbert for editing the manuscript. This work was supported by the Public Health Agency of Canada (D.K., S.M.J. and H.F.), by grants-in-aid for scientific research on priority areas from the Ministries of Education, Culture, Sports, Science, and Technology, Japan (Y.K. and K.S.), by CREST (Japan Science and Technology Agency; Y.K.), and by private grants to Y.K.

Author Information Microarray data were deposited at Arrayexpress with accession number E-TABM-181. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Y.K. (kawaokay@svm.vetmed.wisc.edu).

©2007 Nature Publishing Group

LETTERS

Toxoplasma co-opts host gene expression by injection of a polymorphic kinase homologue

J. P. J. Saeij[1]*, S. Coller[1]*, J. P. Boyle[1], M. E. Jerome[2], M. W. White[2] & J. C. Boothroyd[1]

***Toxoplasma gondii*, an obligate intracellular parasite of the phylum Apicomplexa, can cause severe disease in humans with an immature or suppressed immune system. The outcome of *Toxoplasma* infection is highly dependent on the strain type, as are many of its *in vitro* growth properties[1]. Here we use genetic crosses between type II and III lines to show that strain-specific differences in the modulation of host cell transcription are mediated by a putative protein kinase, ROP16. Upon invasion by the parasite, this polymorphic protein is released from the apical organelles known as rhoptries and injected into the host cell, where it ultimately affects the activation of signal transducer and activator of transcription (STAT) signalling pathways and consequent downstream effects on a key host cytokine, interleukin (IL)-12. Our findings provide a new mechanism for how an intracellular eukaryotic pathogen can interact with its host and reveal important differences in how different *Toxoplasma* lineages have evolved to exploit this interaction.**

Most *Toxoplasma gondii* isolates that have been identified in Europe and North America belong to three distinct clonal lines[2,3], referred to as types I, II and III. The three types differ widely in a number of phenotypes in mice such as virulence, persistence, migratory capacity, attraction of different cell types and induction of cytokine expression[1]. Recent results indicate that such differences might also exist in human infection[4–9]. To test the hypothesis that some of these strain-specific differences are a result of how the strains interact with the host cell, we infected human foreskin fibroblasts (HFFs) with each of the three types and used microarray analysis to investigate differences in host gene expression 24 h later. Significance analysis of microarrays[10] (SAM) identified 105 human complementary DNAs, representing at least 88 unique genes that were regulated in a strain-specific manner (false discovery rate 15%) (Fig. 1a).

If the strain-specific regulation of a host gene has a genetic basis, it should segregate among F1 progeny derived from a cross between two strains that differ in its regulation. We therefore infected HFFs separately with each of 19 F1 progeny derived from crosses between type II and type III parasites and repeated the microarray analyses. The F1 strains formed two distinct clusters and, for a large portion of the genes, the progeny belonging to each cluster modulated human gene expression in either a type II- or a type III-like manner (Fig. 1b).

To identify the loci involved in this differential regulation, we performed a genome-wide scan for association of *Toxoplasma* genetic markers[11] and the expression level of each of the 42,000 human cDNAs on the microarray using the package R/qtl (ref. 12). From this, 3,188 cDNAs, could be mapped to a specific *Toxoplasma* genomic locus (Fig. 2a–c). Interestingly, 1,176 of those cDNAs mapped to chromosome VIIb (Fig. 2b) and, of these, 617 had their highest logarithm of odds (LOD) score around genetic markers L339 and L363 (see Fig. 2c for an example). This indicated that, in the vicinity of these markers, there is at least one polymorphic *Toxoplasma* gene whose product has a strong effect on gene expression in HFFs. This was also corroborated by the fact that all F1 progeny in cluster 1 of Fig. 1b have type III alleles for markers L339 and L363 while all in cluster 2 have type II alleles.

Pathway analysis showed that the group of human genes whose strain-specific modulation mapped to *Toxoplasma* chromosome VIIb was significantly enriched for genes involved in the IL-6, Janus kinase (JAK)/STAT, amyloid processing and IL-4 signalling pathways (Fischer's exact test for all was $P < 5 \times 10^{-3}$; see Supplementary Fig. 2a–e). A broader, network analysis of molecular relationships between genes and gene products resulted in high scores for three networks whose central transcription factors were STAT3 and STAT5b (network 1), JUN (network 2) and hypoxia-inducible factor (HIF)-1A (network 3) (see Supplementary Fig. 2f–h). Given these results and the fact that the IL-4 and IL-6 signalling pathways culminate in the activation of the transcription factors STAT6 and STAT3, respectively[13], we hypothesized that a large part of the strain-specific regulation of host genes is due to differences in how

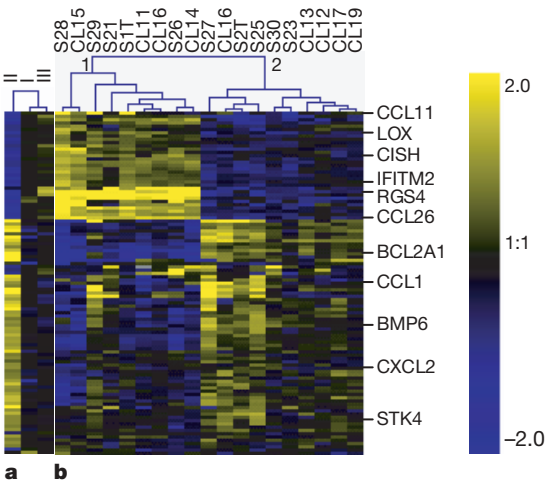


Figure 1 | *Toxoplasma* strain-specific regulation of human gene expression. **a**, HFFs were infected with type I, type II or type III *Toxoplasma* strains. 24 h after infection, expression profiles were obtained using human cDNA arrays. The averaged results (from at least three biological replicates) for median-centred expression levels for cDNAs are displayed using a $\log_2$ blue (low) to yellow (high) scale. Values $\log_2 > 2$ or $\log_2 < -2$ were assigned the values 2 and –2, respectively. **b**, HFFs were infected with 19 F1 progeny derived from crosses (S or CL) between type II and type III strains. Details are as for **a**. Also displayed is the unsupervised clustering of experiments. For a full-size image and array data see Supplementary Fig. 1 and Supplementary data 1.

[1]Department of Microbiology and Immunology, Fairchild Building D305, 300 Pasteur Drive, Stanford University School of Medicine, Stanford, California 94305-5124, USA.
[2]Department of Veterinary Molecular Biology, College of Agriculture, Montana State University, Bozeman, Montana 59717, USA.
*These authors contributed equally to this work.

©2007 Nature Publishing Group

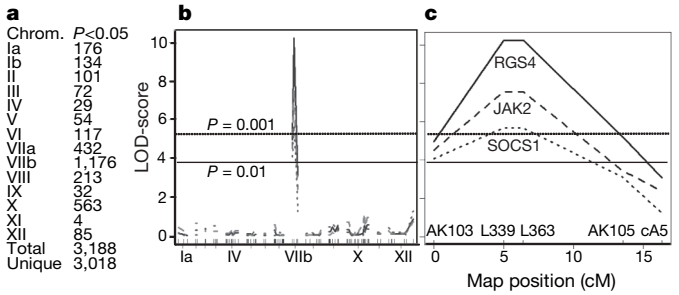


Figure 2 | Genome-wide scans for association of human gene expression with *Toxoplasma* genetic markers. **a**, For each *Toxoplasma* chromosome, the number of cDNAs that mapped significantly ($P < 0.05$, permutation test) to a genetic marker on that chromosome is shown. **b**, Plots indicate the log-likelihood association of expression of the human cDNAs for *RGS4* (solid line), *JAK2* (dashed line) and *SOCS1* (dotted line) with markers aligned across the entire *Toxoplasma* genome in chromosome order. Selected chromosomes are indicated. **c**, An enlargement of Fig. 2b focusing on chromosome VIIb with the names of genetic markers indicated. LOD-score profiles of all significantly mapped human genes are provided as Supplementary data 2.

the different *Toxoplasma* strains intersect the STAT activation pathways, although from the network analysis it is clear that other transcription factors are also probably involved (for example, JUN and HIF1A).

To test whether differences in STAT activation are central to the strain-specific differences, we analysed HFFs infected with type I, II or III parasites using immunofluorescence assays (IFA) and western blotting with antibodies specific for the tyrosine-phosphorylated (activated) forms of STAT3 and STAT6. Both methods revealed that, 18 h after infection, serum-starved HFFs infected with type I or type III strains contain much more activated STAT3 and STAT6 (not shown) than HFFs infected with type II strains (Fig. 3 a, b). At earlier time points (for example, 1 h after infection), HFFs individually infected with any of the three strains showed nuclear localization of tyrosine-phosphorylated STAT3 (data not shown). Somehow, late in infections and with type II parasites only, it seems that phospho-STAT3 Tyr 705 levels in the nucleus drop, presumably because tyrosine phosphorylation is repressed or reversed. In all cases, activation of STAT3 and STAT6 depended on parasite invasion, as uninfected neighbouring cells showed little or no signal. Co-infection of HFFs with type II and type III strains showed strong phosphorylation of STAT3 and STAT6, showing that this is an active property of type I and III strains and that type II strains cannot repress this process (not shown).

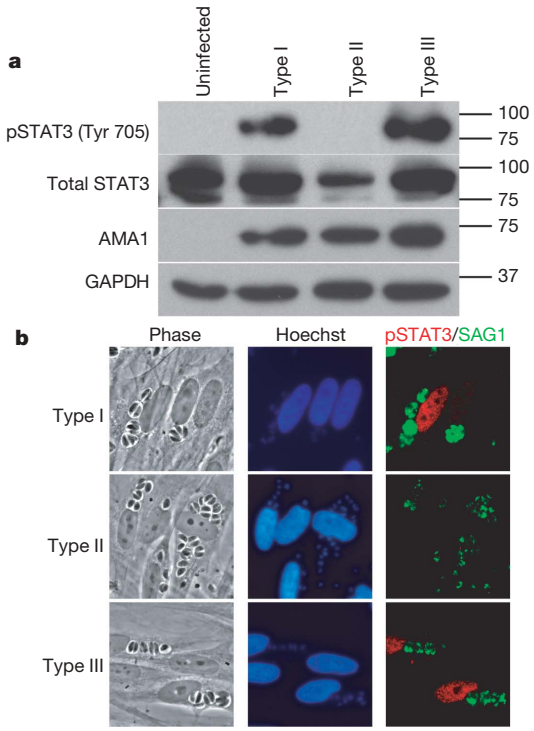


Figure 3 | Strain-specific activation of STAT3. **a**, Western blot analysis of strain-dependent phosphorylation of STAT3. Serum-starved HFFs were infected with a type I, type II or type III *Toxoplasma* strain and 18 h after infection cells were lysed and analysed by immunoblotting. GAPDH (host cell-specific) and AMA1 (parasite-specific) levels are shown as loading controls. **b**, Immunofluorescence analysis of strain-dependent activation of STAT3. HFFs were infected with type I, II or III parasites (MOI 10) for 18 h, fixed and incubated with antibodies against parasite SAG1 or the phosphorylated forms of STAT3 (Tyr 705).

To confirm and refine the preliminary mapping of the *Toxoplasma* genomic region involved in activation of STAT3/6, parasites that had recombinations in the region of interest on chromosome VIIb were genotyped and phenotyped for their capacity to activate STAT3/6. These data allowed us to narrow down the locus involved to a region between L363 and AK104, representing a maximum size of 0.56 Mb (see Supplementary Fig. 3). This interval on VIIb contains 78 predicted proteins (http://www.toxodb.org). Assuming that the crucial gene encodes a polymorphic, secreted protein that can substantially alter host cell transcription, we identified *ROP16* as the most likely candidate. First, there are 57 single-nucleotide polymorphisms (SNPs; 34 of which are non-synonymous) between the type II allele of *ROP16* and the type I and III alleles, which are virtually identical (5 SNPs), consistent with the genetic evidence that the type II allele functions completely differently from those of types I and III (see Supplementary Fig. 4). Second, ROP16 has high homology to serine/threonine protein kinases, which are potent modulators of many cell functions. And, third, ROP16 localizes to the *Toxoplasma* rhoptries[14], a set of apical organelles that can secrete vesicle-like bodies (termed 'evacuoles') into the host cell upon invasion[15].

To characterize ROP16 further, a type I strain was engineered to express a carboxy-terminally haemagglutinin (HA)-tagged version of the protein derived from either a type I or type II background; the resulting strains were designated type I:ROP16$_{I}$ and type I:ROP16$_{II}$, respectively. IFA showed that the tagged ROP16 was correctly targeted to the rhoptries and that within 10 min of invasion it had entered the host cell nucleus; it was still visible at 24 h but the brightest relative staining occurred between 10 min and 4 h (Fig. 4) with an apparent diminution of signal thereafter. Similar results were obtained when the mouse macrophage cell line RAW 264.7 was used as the host cell (not shown). Localization of HA-tagged ROP16 to the host cell nucleus depended on the putative nuclear localization signal

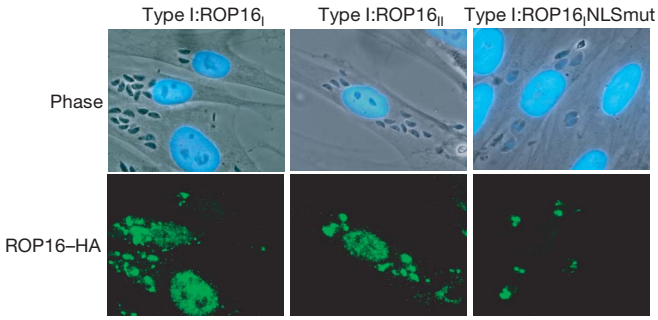


Figure 4 | ROP16 is secreted from the parasite and localizes to the host cell nucleus. Type I parasites expressing HA-tagged type I ROP16 (Type I:ROP16$_{I}$, left), type II ROP16 (Type I:ROP16$_{II}$, middle) or type I ROP16 with a mutated NLS (Type I:ROP16$_{I}$NLSmut, right) were added (MOI 10) to HFFs and cells were fixed after 4 h. ROP16 was visualized using anti-HA antibodies followed by incubation with secondary (anti-rat Alexa 488) antibodies.

©2007 Nature Publishing Group

(NLS) in ROP16 (RKRKRK at residues 339–344, inclusive): mutation of lysine residues 340, 342 and 344 to methionines (yielding RMRMRM) markedly diminished nuclear localization (Fig. 4).

In both HFFs and macrophages, the intensity of ROP16 staining in the nucleus depended on the number of invasion events (as determined by the number of parasitophorous vacuoles), not the total number of parasites present, which increases as a function of time (Fig. 4 and data not shown). Combined with the fact that ROP16 reaches the nucleus within 10 min of initial contact, these data indicate that ROP16 is introduced during the invasion process and not in a slow, steady release from intracellular parasites. Recently, we have seen that a second rhoptry protein, in this case a protein phosphatase, is injected into host cells with similar kinetics to ROP16. It also possesses an NLS and reaches the host nucleus with similar kinetics[16].

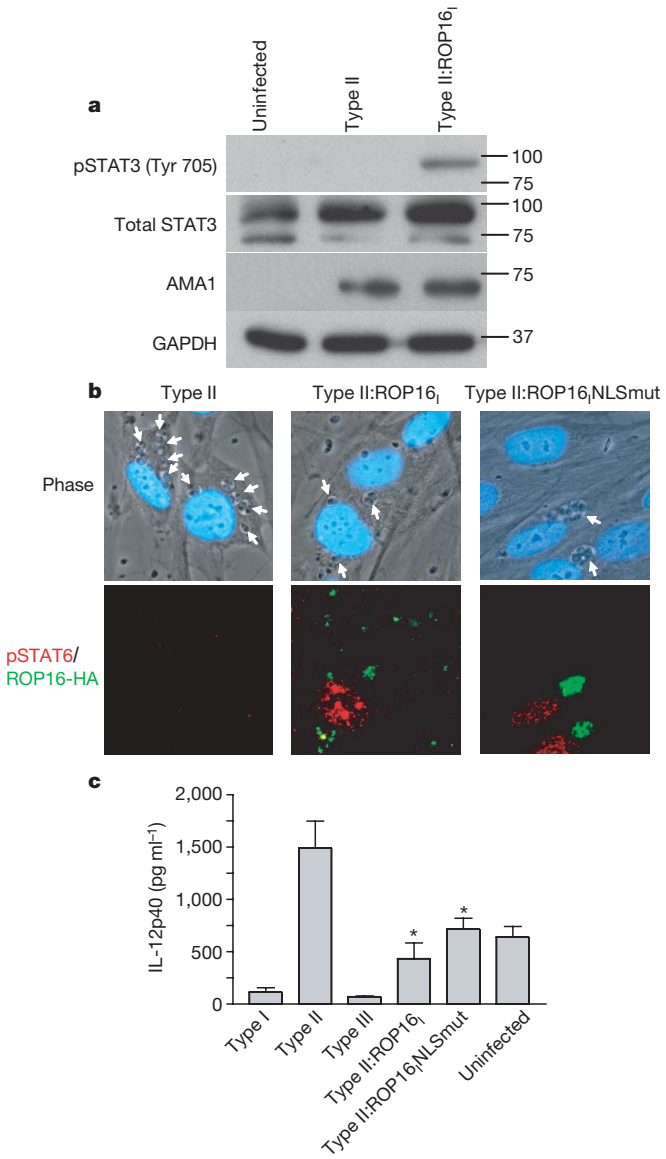


Figure 5 | ROP16 mediates strain-specific activation of STAT3/6 and consequent downstream effects on IL-12. a, Type II parasites expressing the type I ROP16 gene induce phosphorylation of STAT3. HFFs were infected with type II or type II:ROP16$_I$ and lysates were prepared for immunoblot analysis as described in the legend to Fig. 3a. **b**, Type II parasites expressing the normal type I ROP16 gene or the type I ROP16 NLS mutant gene induce translocation of phosphorylated STAT6 to the host cell nucleus. White arrows indicate parasites. **c**, Type II parasites expressing the type I ROP16 gene or the type I ROP16 NLS mutant gene inhibit the production of IL-12p40 from infected macrophages. Error bars represent s.e.m. of three replicates. Data shown are from one representative experiment out of three. Asterisks denote significantly less IL-12p40 production relative to type II parasites (Student's *t*-test).

To test whether ROP16 is involved in the strain-specific activation of STAT3/6, we engineered type II strains to express an additional, type I allele of ROP16 (designated type II:ROP16$_I$). Western blot and IFA revealed that, as before, HFFs infected with the parental type II strain showed little if any STAT3/6 activation 18 h after infection, whereas those infected with the type II:ROP16$_I$ showed strong activation in both assays (Fig. 5a, b). These experiments show that ROP16 has an important role in maintaining the activation of STAT3 and STAT6 in infected HFFs and indicate that the type I allele is probably dominant over the type II version.

Tyrosine phosphorylation of STAT3 and STAT6 generally occurs in the cytoplasm of a cell. It is therefore surprising that ROP16 seems to be concentrated in the host cell nucleus. To investigate whether nuclear localization is essential for its function, we engineered a type II strain to express a type I allele of ROP16 carrying the NLS disruption described above (RKRKRK to RMRMRM). Infection with the resulting strain, designated type II:ROP16$_I$NLSmut, showed a comparable level of STAT6 activation as infection with the type II:ROP16$_I$ strain (Fig. 5b). These results indicate that ROP16 can probably execute its function in the cytosol (perhaps as it transits to the nucleus) although we cannot exclude the possibility that the small amount of ROP16 that still goes to the nucleus in the NLS mutant (not shown) is responsible for its function in cells infected with this strain.

To investigate how ROP16 mediates its effect on STAT3/6, we used a battery of antibodies that recognize the activated (phosphorylated) forms of proteins upstream of the STAT3/6 signalling pathways. We found no strain differences in the activation of extracellular signal regulated kinase (ERK)1/2, P38, Jun amino-terminal kinase (JNK), AKT, JAK2 and tyrosine kinase (TYK)2. Combined with the fact that all three strains activate STAT3/6 early (see above and Supplementary Fig. 5), these results indicate that the pathway that leads to the initial activation of these key transcription factors is similar between strains but that HFFs infected with type I/III strains are actively prevented from subsequent downregulation.

Our results provide a potential molecular basis for at least one of the strain-dependent differences in how a host cell responds to *Toxoplasma* infection: type II but not type I or III strains induce mouse macrophages to produce high levels of IL-12 (both p40 and p70; ref. 17) and constitutive activation of STAT3 by type I parasites prevents lipopolysaccharide-triggered production of IL-12p40 by macrophages[18]. To determine whether ROP16 is involved in these differences, we examined STAT3 phosphorylation and IL-12p40 production by murine RAW264.7 macrophages infected with our engineered parasites and stimulated with interferon (IFN)-γ and LPS. The amount of phosphorylated STAT3 in macrophages infected with the different *Toxoplasma* strains paralleled the results described above for HFFs (see Supplementary Fig. 6). We confirmed that infection of macrophages with type II, but not with the type I or III strains, results in very high secretion of IL-12p40 and further showed that the type II:ROP16$_I$ and type II:ROP16$_I$NLSmut strains elicit significantly less IL-12p40 than parental type II parasites (Fig. 5c), as predicted by our results showing increased activation of STAT3 by these strains. Such effects could be an important part of the different disease outcomes seen with the type I, II and III strains. The facts that the type II strain induced more IL-12p40 than IFN-γ and LPS alone, and that the type II:ROP16$_I$ strain did not bring the IL-12p40 down to the levels of type I and III strains, indicate that in addition to ROP16, other molecules might be involved in the strain-specific regulation of IL-12p40.

To come full circle and show that because of or in addition to its role in strain-specific activation of STAT3/6, ROP16 is involved in the strain-specific differences originally observed in the human cDNA microarray experiments, we compared the gene expression profiles of HFFs infected with type II and type II:ROP16$_I$ strains.

©2007 Nature Publishing Group

Of 13,382 cDNAs that met the quality control threshold, 127 were previously shown (Fig. 2a) to map to chromosome VIIb with an average difference greater than twofold between the original F1 progeny based on which ROP16 allele they carried. Of these 127 cDNAs, 29 (out of 242 total) were differentially regulated (average of at least twofold difference in the two arrays) when comparing the type II and type II:ROP16$_I$ strains. This is significantly more genes than expected by chance (hypergeometric distribution, $P < 10^{-15}$), confirming a role for ROP16 in this strain-specific regulation of host gene expression. Furthermore, all 29 genes changed in the expected direction (See Supplementary data 3).

The phenomenology of ROP16 introduction is analogous to type III and type IV secretion systems in prokaryotes[19] but there are no indications for the existence of such systems outside bacteria. One of the earliest events in *Toxoplasma* invasion is a break in the host cell membrane[20] and this might represent the moment when injection of ROP16 occurs, although we have been unable to detect ROP16 free in the cytosol or even associated with evacuoles[15] (data not shown). Our results do not reveal the primary target of ROP16. It probably does not directly phosphorylate STAT3/6 but instead is likely to intersect pathways that lead to or maintain such activation. This could involve something akin to what has been reported[21] for v-Abl, an oncogenic tyrosine kinase that causes constitutive activation of Jaks and STATs by disrupting suppressor of cytokine signalling (SOCS)-1 function (SOCS-1 normally targets Jaks and STATs for proteasomal degradation). Infection with type I and III strains results in a higher level of SOCS-1, SOCS-2 and SOCS-3 mRNA than infection with type II (see Supplementary data 2) and this difference maps to ROP16 (Fig. 2). This result is consistent with the fact that STAT3 is a positive regulator of SOCS (that is, the sustained activation of STAT3 results in their upregulation) and indicates that the mechanism involved in the type I and III-specific maintenance of activated STAT3 involves a block in SOCS function or some other downstream effect rather than a decrease in SOCS expression.

Rhoptries are a defining property of all Apicomplexa and some of the proteins found in the *Toxoplasma* rhoptries have homologues in *Plasmodium*. Thus, the ability to co-opt host cell gene expression for the parasite's own purposes might be found in other Apicomplexa that also are otherwise trapped within a membrane-limited vacuole (for example, *Plasmodium* species as they grow within hepatocytes). This represents a previously undescribed mechanism for interaction between eukaryotic pathogens and their hosts.

METHODS

Details of parasite strains, genetic crosses and microarray analysis can be found in Supplementary Information.

Immunofluorescence analysis. Parasites were allowed to invade confluent HFF monolayers on coverslips for 0.17, 0.5, 1, 4, 6, 8, 18 or 24 h. The cells were then fixed, blocked and permeabilized. Coverslips were incubated with 3F10 (anti-HA) antibody, antibodies specific for the tyrosine phosphorylated forms of STAT3 and STAT6, or a mouse monoclonal antibody against surface antigen (SAG)-1 (DG52)[22]. Fluorescent secondary antibodies and Hoechst dye were used for antigen and DNA visualization, respectively.

Western blots. Parasites (MOI 10) were added to HFFs and infection was allowed to proceed for 18 h. Western blots were performed as described[23] using antibodies specific for total STAT3 (Cell Signaling Technologies), GAPDH (Calbiochem), AMA1 (mAb B3.90[24]) or for the phosphorylated forms of STAT3 (phospho-Tyr705) and STAT6 (phospho-Tyr641) (both from Cell Signaling Technologies).

IL-12 enzyme-linked immunosorbent assay (ELISAs). RAW264.7 macrophages were activated with IFN-γ (20 U ml^{-1}) and LPS (100 ng ml^{-1}) and 1 h later parasites were added at a multiplicity of infection (MOI) of 10. After 24 h of infection, supernatants were collected and IL-12p40 ELISAs were performed as per manufacturer's recommendations (BD Biosciences).

Received 13 July; accepted 1 November 2006.
Published online 20 December 2006.

1. Saeij, J. P., Boyle, J. P. & Boothroyd, J. C. Differences among the three major strains of *Toxoplasma gondii* and their specific interactions with the infected host. *Trends Parasitol.* **21,** 476–481 (2005).
2. Howe, D. K. & Sibley, L. D. *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *J. Infect. Dis.* **172,** 1561–1566 (1995).
3. Darde, M. L., Bouteille, B. & Pestre-Alexandre, M. Isoenzyme analysis of 35 *Toxoplasma gondii* isolates and the biological and epidemiological implications. *J. Parasitol.* **78,** 786–794 (1992).
4. Grigg, M. E., Ganatra, J., Boothroyd, J. C. & Margolis, T. P. Unusual abundance of atypical strains associated with human ocular toxoplasmosis. *J. Infect. Dis.* **184,** 633–639 (2001).
5. Howe, D. K., Honore, S., Derouin, F. & Sibley, L. D. Determination of genotypes of *Toxoplasma gondii* strains isolated from patients with toxoplasmosis. *J. Clin. Microbiol.* **35,** 1411–1414 (1997).
6. Fuentes, I., Rubio, J. M., Ramirez, C. & Alvar, J. Genotypic characterization of *Toxoplasma gondii* strains associated with human toxoplasmosis in Spain: direct analysis from clinical samples. *J. Clin. Microbiol.* **39,** 1566–1570 (2001).
7. Khan, A. *et al.* Genotyping of *Toxoplasma gondii* strains from immunocompromised patients reveals high prevalence of type I strains. *J. Clin. Microbiol.* **43,** 5881–5887 (2005).
8. Darde, M. L., Villena, I., Pinon, J. M. & Beguinot, I. Severe toxoplasmosis caused by a *Toxoplasma gondii* strain with a new isoenzyme type acquired in French Guyana. *J. Clin. Microbiol.* **36,** 324 (1998).
9. Carme, B. *et al.* Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana. *J. Clin. Microbiol.* **40,** 4037–4044 (2002).
10. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA* **98,** 5116–5121 (2001).
11. Khan, A. *et al.* Composite genome map and recombination parameters derived from three archetypal lineages of *Toxoplasma gondii*. *Nucleic Acids Res.* **33,** 2980–2992 (2005).
12. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19,** 889–890 (2003).
13. Ihle, J. N. The Stat family in cytokine signaling. *Curr. Opin. Cell Biol.* **13,** 211–217 (2001).
14. Bradley, P. J. *et al.* Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in *Toxoplasma gondii*. *J. Biol. Chem.* **280,** 34245–34258 (2005).
15. Hakansson, S., Charron, A. J. & Sibley, L. D. *Toxoplasma* evacuoles: a two-step process of secretion and fusion forms the parasitophorous vacuole. *EMBO J.* **20,** 3132–3144 (2001).
16. Gilbert, L. A., Ravindran, S., Turetzky, J. M., Boothroyd, J. C. & Bradley, P. J. *Toxoplasma gondii* targets a protein phosphatase 2C to the nucleus of infected host cells. *Eukaryot. Cell* (published online 3 November 2006; doi:10.1128/EC/00309-06).
17. Robben, P. M. *et al.* Production of IL-12 by macrophages infected with *Toxoplasma gondii* depends on the parasite genotype. *J. Immunol.* **172,** 3686–3694 (2004).
18. Butcher, B. A. *et al.* IL-10-independent STAT3 activation by *Toxoplasma gondii* mediates suppression of IL-12 and TNF-alpha in host macrophages. *J. Immunol.* **174,** 3148–3152 (2005).
19. Remaut, H. & Waksman, G. Structural biology of bacterial pathogenesis. *Curr. Opin. Struct. Biol.* **14,** 161–170 (2004).
20. Suss-Toby, E., Zimmerberg, J. & Ward, G. E. *Toxoplasma* invasion: The parasitophorous vacuole is formed from host cell pasma membrane and pinches off via a fission pore. *Proc. Natl Acad. Sci. USA* **93,** 8413–8418 (1996).
21. Limnander, A. *et al.* v-Abl signaling disrupts SOCS-1 function in transformed pre-B cells. *Mol. Cell* **15,** 329–341 (2004).
22. Burg, J. L., Perelman, D., Kasper, L. H., Ware, P. L. & Boothroyd, J. C. Molecular analysis of the gene encoding the major surface antigen of *Toxoplasma gondii*. *J. Immunol.* **141,** 3584–3591 (1988).
23. Chan, S. M. *et al.* Protein microarrays for multiplex analysis of signal transduction pathways. *Nature Med.* **10,** 1390–1396 (2004).
24. Mital, J. *et al.* Conditional expression of *Toxoplasma gondii* apical membrane antigen-1 (TgAMA1) demonstrates that TgAMA1 plays a critical role in host cell invasion. *Mol. Biol. Cell* **16,** 4341–4349 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by grants from the NIH (to J.C.B., J.P.B. and M.W.W.), the Ellison Medical Foundation (to J.C.B.), the USDS (to M.W.W.) and the California Universitywide AIDS Research Program (to J.P.J.S. and S.C.). We thank K. Broman for help with R/qtl and the Stanford Functional Genomics Facility for the human cDNA microarrays used for this study.

Author Contributions J.P.J.S. and S.C. contributed equally to this work. J.P.J.S. performed the microarrays and pathway analyses. S.C. and J.P.J.S. performed the experiments in Fig. 3. S.C. performed the experiments in Fig. 4 and Fig 5. J.P.B., M.E.J. and M.W.W. performed the genetic crosses that produced the progeny D3X1 and JD4. J.P.B. genotyped D3X1 and JD4. J.P.J.S., S.C., J.P.B. and J.C.B. wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information The microarray data have been deposited in ArrayExpress with the accession number E-MEXP-783. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.B. (john.boothroyd@stanford.edu).

©2007 Nature Publishing Group

LETTERS

CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast

Seiji Tanaka[1,2,3], Toshiko Umemori[1,3], Kazuyuki Hirai[1,3], Sachiko Muramatsu[1], Yoichiro Kamimura[1,2,3]† & Hiroyuki Araki[1,2,3]

In eukaryotic cells, cyclin-dependent kinases (CDKs) have an important involvement at various points in the cell cycle. At the onset of S phase, active CDK is essential for chromosomal DNA replication[1], although its precise role is unknown. In budding yeast (*Saccharomyces cerevisiae*), the replication protein Sld2 (ref. 2) is an essential CDK substrate[3], but its phospho-mimetic form (Sld2-11D) alone neither affects cell growth[4] nor promotes DNA replication in the absence of CDK activity[4], suggesting that other essential CDK substrates promote DNA replication. Here we show that both an allele of *CDC45* (*JET1*) and high-copy *DPB11*, in combination with Sld2-11D, separately confer CDK-independent DNA replication. Although Cdc45 is not an essential CDK substrate, CDK-dependent phosphorylation of Sld3, which associates with Cdc45 (ref. 5), is essential and generates a binding site for Dpb11. Both the *JET1* mutation and high-copy *DPB11* by-pass the requirement for Sld3 phosphorylation in DNA replication. Because phosphorylated Sld2 binds to the carboxy-terminal pair of BRCT domains in Dpb11 (ref. 4), we propose that Dpb11 connects phosphorylated Sld2 and Sld3 to facilitate interactions between replication proteins, such as Cdc45 and GINS. Our results demonstrate that CDKs regulate interactions between BRCT-domain-containing replication proteins and other phosphorylated proteins for the initiation of chromosomal DNA replication; similar regulation may take place in higher eukaryotes.

Pre-replicative complexes form on replication origins from late M phase to G1 phase, when CDK activity is low[6–8]. Many replication proteins, including replicative DNA polymerases, assemble on pre-replicative-complex-containing origins upon the action of Dbf4-dependent kinase (DDK or Cdc7) and CDK[9–12] (Supplementary Fig. 1). Replication proteins Sld3 and Cdc45 form a complex and associate with early firing origins in the absence of these kinase activities[5,13]. Sld3 protein has 12 CDK phosphorylation motifs[5]. A newly constructed *sld3-12A* allele, where alanine replaced serine and threonine residues at all 12 CDK phosphorylation motifs, could not substitute for wild-type function (Fig. 1a). By combining various alanine substitutions, a mutant bearing simultaneous alanine substitutions for Thr 600 and Ser 622 (*sld3-2A*) appeared to be lethal (Fig. 1a). An *sld3* mutant that was unmodified at either Thr 600 or Ser 622 but contained substitutions at the remaining 11 phosphorylation motifs (11A-T600 or 11A-S622) was viable. *Sld3-5td* cells (the *sld3-5td* allele encodes a temperature-dependent degron with the *sld3-5* mutation[5]) harbouring *sld3-2A* or *sld3-12A* alleles were released from G1 arrest after inactivation of Sld3-5td but did not increase DNA content for 120 min, and then gradually approached to 2C content (Fig. 1b), indicating that they are defective in DNA replication. Moreover, recombinant Cdc28-Clb5 kinase[4] phosphorylated Thr 600 and Ser 622 of the purified glutathione *S*-transferase (GST)–Sld3-C3 protein *in vitro* (Supplementary Fig. 2b). Thr 600 and Ser 622 were phosphorylated *in vivo* at the G1/S boundary and S phase with the same timing as Orc6 phosphorylation (Fig. 1c). Orc6 is phosphorylated by S-phase and M-phase CDKs (S-CDKs and M-CDKs) and migrates slower than the hypo-phosphorylated form on SDS–PAGE[14]. Furthermore, neither site of Sld3 (Thr 600 and Ser 622) was phosphorylated in cells arrested at G1 phase by Sic1ΔNT, an inhibitor of S- and M-CDKs (Fig. 1d): phosphorylation of both residues was reduced in cells lacking the S-phase cyclins (S-cyclins) Clb5 and Clb6. In these cells, M-cyclin compensates for S-cyclins, and chromosome DNA replication is delayed[15] (Supplementary Fig. 2c). These results suggest that S-CDKs phosphorylate Thr 600 and Ser 622 primarily, both *in vitro* and *in vivo*, and this phosphorylation is essential for chromosomal DNA replication.

The *sld3-A600* mutant, which bears an alanine substitution for Thr 600 but not for Ser 622 (Fig. 1a), grew poorly on a plate containing hydroxyurea (an inhibitor of DNA replication) at 25 °C, and even on a YPDA plate at 16 °C. *DPB11* on a high-copy plasmid restored the growth of these cells (Supplementary Fig. 3a). Moreover, in the presence of high-copy *DPB11*, *sld3-2A* could replace wild-type *SLD3* (Supplementary Fig. 3b), indicating that high-copy *DPB11* by-passes the CDK requirement for phosphorylation of Sld3. Because reduced interaction is often overcome by an increased dosage of partner proteins, it seems likely that the reduced interaction between Sld3-2A and Dpb11 conferred a growth defect that could be rescued by increased dosage of Dpb11. As expected, Dpb11 interacted with Sld3 but not Sld3-2A (Fig. 2a, b), suggesting that the interaction is phosphorylation-dependent. Dpb11 has two pairs of tandem BRCT domains[16]. The tandem BRCT domains are known as a phosphopeptide binding domain[17], and the C-terminal pair of BRCT domains in Dpb11 binds to phosphorylated Sld2 (ref. 4). The amino-terminal pair (Dpb11-N), but not the C-terminal pair (Dpb11-C), appeared to interact with Sld3 (Fig. 2b). The shorter C-terminal fragment, Sld3-C3, also interacted with Dpb11-N and this interaction was also dependent on Thr 600 and Ser 622 (Fig. 2b). Furthermore, purified Sld3-C3 bound to purified Dpb11 and this binding was markedly increased by treatment of Sld3-C3 with recombinant Cdc28-Clb5 (Fig. 2c). These results suggest that Dpb11 binds to the C-terminal portion of Sld3 in a manner facilitated by CDK-catalysed phosphorylation of Thr 600 and Ser 622. Because non-phosphorylated Sld3 (Sld3-2A and Sld3-12A) is defective in DNA replication (Fig. 1b), the interaction between Dpb11 and phosphorylated Sld3 is suggested to be essential for DNA replication. This is supported by the finding that fusion of Sld3 and the C-terminal pair of BRCT domains in Dpb11 by-passes the requirement for Sld3 phosphorylation[18].

[1]Division of Microbial Genetics, National Institute of Genetics, Research Organization of Information and Systems, and [2]Department of Genetics, SOKENDAI, Yata 1111 Mishima, Shizuoka 411-8540, Japan. [3]CREST, Kawaguchi, Saitama 332-0012, Japan. †Present address: Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA.

©2007 Nature Publishing Group

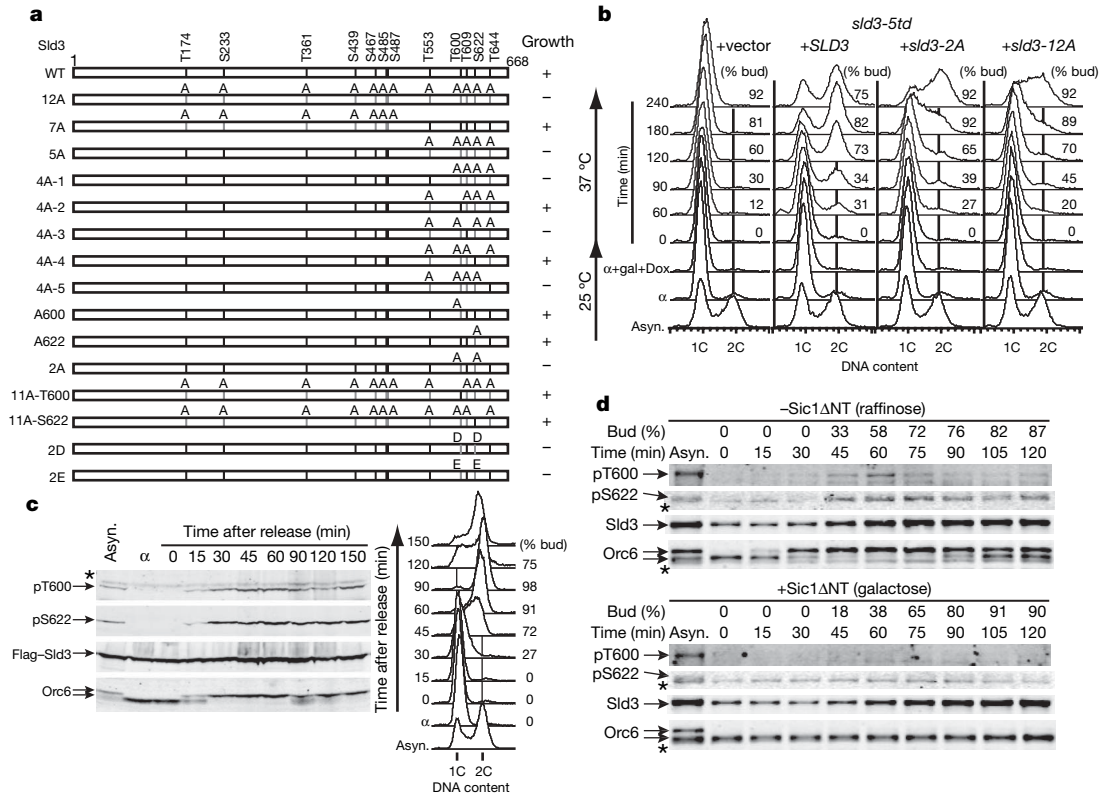


Figure 1 | CDK-catalysed essential phosphorylation of the Sld3 protein. **a**, Various Sld3 alanine substitution constructs tested for cell growth in a plasmid-shuffling assay. Vertical lines indicate serine and threonine residues at CDK phosphorylation motifs. A, D and E indicate alanine, aspartate and glutamate substitution, respectively. **b**, Flow cytometry of synchronized YTU8 (*sld3-5td*) cells harbouring the indicated *SLD3* allele. Cells were arrested in G1 (α) and released after degradation of Sld3-5td (α+gal+Dox). Budding indices are also shown. Asyn, asynchronous cells. **c**, Protein samples from synchronized YTU9 cells were analysed with antibodies against phospho-Thr 600, phospho-Ser 622, Flag and Orc6, respectively. Asterisk, nonspecific signal. **d**, Protein samples from synchronized YST21-1 (*GALp-SIC1ΔNT*) cells were analysed as in **c**.

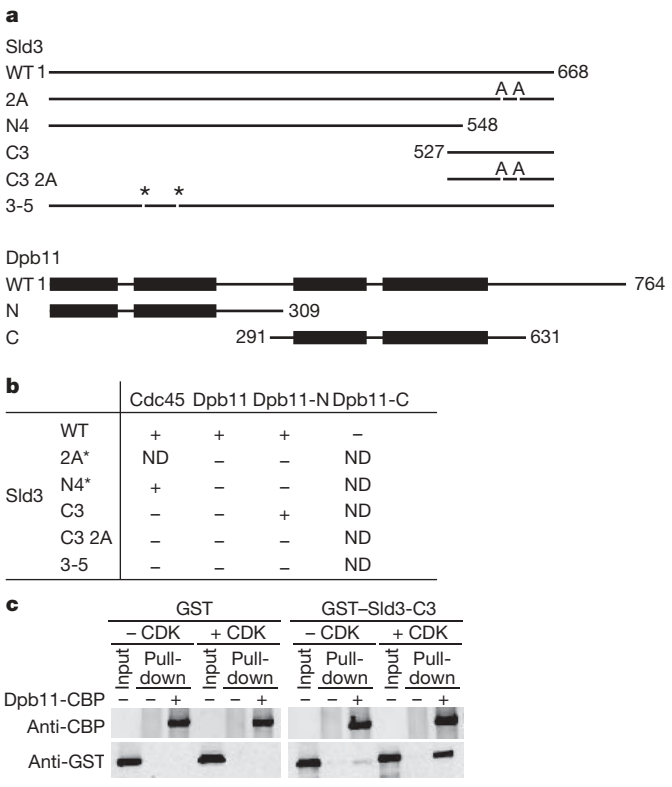


Figure 2 | Phosphorylation-dependent interaction between Dpb11 and Sld3. **a**, Sld3 and Dpb11 constructs used in the two-hybrid assay. Sld3 and its derivatives were constructed on activation domain vector. Numbers represent amino acid position. Asterisks on 3-5 represent the *sld3-5* mutation sites[5]. Thick black bars on Dpb11 represent BRCT domains. **b**, Summary of the two-hybrid results. ND, not determined. The asterisk indicates that simultaneous introduction of high-copy *CDC45* is required for cell viability probably because Sld3-2A and N4 form inactive complexes with Cdc45. **c**, Phosphorylation-dependent complex formation between Sld3 and Dpb11 *in vitro*. Purified GST or GST–Sld3-C3 (see Fig. 2a and Methods) pre-treated with CDK (+CDK) or buffer (−CDK) was mixed with CBP-tagged Dpb11 pre-bound to protein A beads.

In parallel, we isolated a dominant *JET1* mutation of *CDC45* (jumping CDK essentiality with Sld2-11D to initiate DNA replication; for brevity, we use *JET1* hereafter) (see Supplementary Fig. 4 for details). A single copy of the *JET1* gene suppressed the growth defect caused by the *sld3-2A* mutation (Supplementary Fig. 5a) but not the N-terminal *sld3-5* mutation[5] (data not shown; see Fig. 2a for mutation sites), indicating that the *JET1* mutation, like high-copy *DPB11*, by-passes the requirement for phosphorylation at Thr 600 and Ser 622. Because Cdc45 interacted with the N-terminal portion of Sld3 in a two-hybrid assay (Fig. 2a, b) and Cdc45/Jet1 and Sld3 proteins expressed in *Escherichia coli* bind directly (Supplementary Fig. 6), it is further suggested that Cdc45/Jet1 affects the interaction between Dpb11 and Sld3. When the interaction between Cdc45 and Sld3 was reduced by the *sld3-5* mutation[5], the interaction between Sld3 and Dpb11-N diminished (Fig. 2a, b); introduction of high-copy *CDC45* partially restored this interaction (Supplementary Fig. 5b). In addition, introduction of *JET1*, but not wild-type *CDC45*, into the two-hybrid assay partially restored the interaction between

©2007 Nature Publishing Group

Sld3-2A and Dpb11 (Supplementary Fig. 5c). *JET1* also suppressed the *dpb11-24* and *dpb11-26* mutations in the N-terminal BRCT domains of Dpb11 that bind to Sld3, but not the *dpb11-31* and *dpb11-1* mutations in the C-terminal BRCT domains (Supplementary Fig. 5d). These results suggest that Jet1 binds to Sld3 and then facilitates the interaction between Sld3 and Dpb11 without phosphorylation.

We have shown that Sld2 and Sld3 are essential CDK targets for DNA replication. If simultaneous phosphorylation of Sld2 and Sld3 is a minimal requirement to promote DNA replication, by-pass of these phosphorylations might promote DNA replication in the absence of CDK activity; indeed, this is the case. Introduction of either high-copy *DPB11* or *JET1* together with Sld2-11D (the phospho-mimetic form of Sld2) led to an increase in the DNA content of cells arrested in G1 with α-factor (Fig. 3a, b). This increase in DNA content depended on the production of Sld2-11D, and especially on an aspartate substitution for Thr 84 in Sld2 (Supplementary Fig. 7b). Thr 84 is an essential phosphorylation site for binding to Dpb11 (ref. 4). Furthermore, the increase of DNA content in *JET1* cells does not require CDK activity, because DNA content increased in *JET1* cells arrested by Sic1ΔNT[19], a potential inhibitor of S- and M-CDKs, and 1-NM-PP1[20], an inhibitor of all the CDKs in *cdc28-as1* cells (Fig. 3c, d; see also the legend to Supplementary Fig. 1 for details).

Several findings indicate that the increase in DNA content was caused by chromosomal DNA replication in *JET1* cells. First, an increase in DNA content depends on replication factors Cdc6 and DDK. The initiation of DNA replication requires Cdc6-dependent formation of pre-replicative complexes on replication origins. DNA content in cells expressing *CDC6* increased, whereas it did not increase in the absence of *CDC6* (Supplementary Fig. 8a). DDK (Cdc7-Dbf4)[21] is also required for the initiation of DNA replication. DNA content in *JET1* cells arrested at G1 (which displays the lowest DDK activity) increased significantly compared to wild-type cells, although this was slower compared with cells expressing *DBF4* from the tetracycline promoter (Fig. 3b; see also Supplementary Fig. 7b). Repressing *DBF4* transcription, or heat-inactivating Cdc7 using the *cdc7-4* temperature-sensitive mutation, abolished the increase of DNA content (Fig. 4a; see also Supplementary Fig. 8b). Second, normal replication forks are established in *JET1* cells. A complex containing Mcm, Cdc45 and GINS has been identified in *Xenopus* egg extracts[22], budding yeast[23] and *Drosophila*[24] and has been suggested to associate with replication forks. When DNA replication was induced by expression of *SLD2-11D* and the *JET1* mutation in the presence of Sic1ΔNT, the mutant Cdc45 protein (Jet1) co-precipitated with Mcm2 and with subunits of GINS, Sld5 and Psf3 (Fig. 4b), suggesting that a complex containing Mcm, Cdc45 and GINS formed. Third, chromosomal DNA is synthesized in a semi-conservative manner. When Sld2-11D was induced in α-factor-arrested *JET1* cells, chromosomal DNA pre-labelled with heavy isotopes (HH) moved to hybrid fractions (HL) and later to light fractions (LL) (Fig. 4c; see also

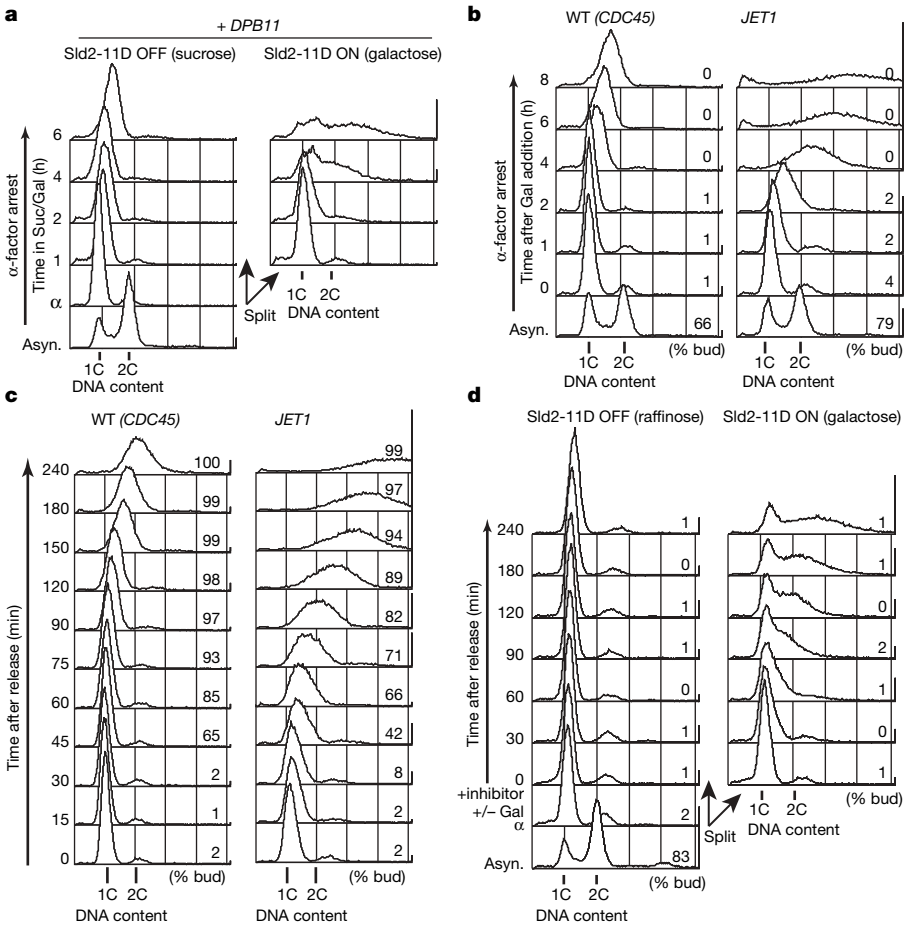


Figure 3 | CDK-independent increase of cellular DNA content. **a**, Flow cytometry of YST516 (*GALp-sld2-11D TetO$_2$-DBF4*) cells harbouring high-copy *DPB11* arrested with α-factor with or without induction of Sld2-11D. Control data of cells with vector are shown in Supplementary Fig. 7a. **b**, Flow cytometry of *JET1* cells. YST516 (wild type: *GALp-sld2-11D TetO$_2$-DBF4*) and YST528 (*JET1 GALp-sld2-11D TetO$_2$-DBF4*) cells were treated as in **a**. Budding indices are indicated (% bud). **c**, Flow cytometry of cells arrested in late G1 with Sic1ΔNT, an inhibitor of S- and M-CDKs. YST535 (WT (*CDC45*)) and YST578 (*JET1*) cells were used with induction of Sld2-11D. Levels of Sld2-11D, Orc6 and Sic1ΔNT proteins were also monitored (Supplementary Fig. 7c). **d**, Flow cytometry of YST604 (*JET1 GALp-sld2-11D TetO$_2$-DBF4 cdc28-as1*) cells arrested with CDK inhibitor, 1-NM-PP1, with or without induction of Sld2-11D. Levels of Sld2-11D and Orc6 proteins were also monitored (Supplementary Fig. 7d).

©2007 Nature Publishing Group

Supplementary Fig. 8d, e). This indicates that DNA is synthesized in a semi-conservative manner more than once. Because CDK only permits cells to replicate chromosomal DNA once per cell cycle[1,6–9,14], the result suggests that in the absence of CDK activity, *JET1* cells can over-ride this restraint of the cell cycle in combination with Sld2-11D. This re-initiation seems to be lethal, because the cells inducing DNA replication by expression of *SLD2-11D* and *JET1* lost viability (data not shown). Fourth, DNA synthesis starts from the same origins as does normal replication. Among the origins (ARS305, ARS306 and ARS501) we examined, bubble arcs, indicative of initiation, were detected from the strong early-firing origins ARS305 and ARS306, in two-dimensional gel assays[25] (Fig. 4d).

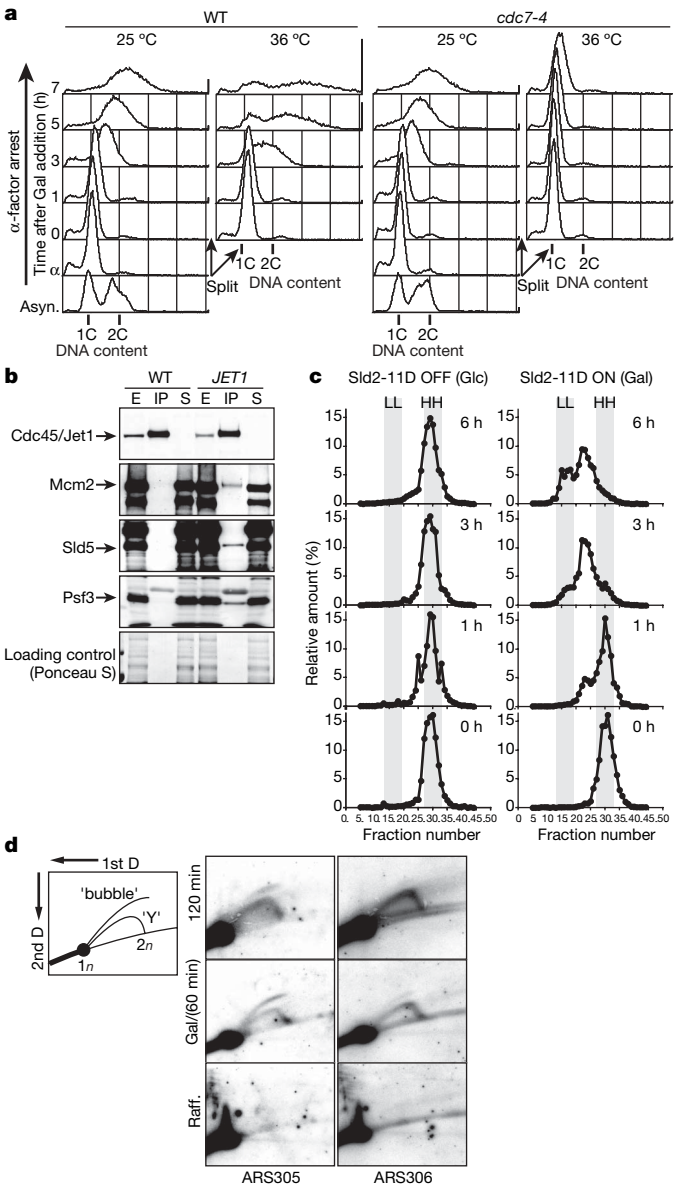


Figure 4 | CDK-independent chromosome DNA replication. a, Flow cytometry of YST630 (WT: *JET1*) and YST631 (*cdc7-4*: *cdc7-4 JET1*) cells arrested with α-factor with induction of Sld2-11D at 25 °C and 36 °C. **b**, CDK-independent co-precipitation of Jet1, Mcm and GINS. Cdc45/Jet1–3Flag was precipitated with anti-Flag beads from extracts of YST674 (WT: *CDC45–3FLAG1HA GALp-sld2-11D GALp-DBF4 GALp-SIC1ΔNT*) and YST678 (*JET1*: *JET1–3FLAG1HA GALp-sld2-11D GALp-DBF4 GALp-SIC1ΔNT*) cells arrested with Sic1ΔNT with induction of Sld2-11D. Cdc45/Jet1–3Flag1HA, Mcm2, Sld5 and Psf3 were detected using anti-Flag, anti-Mcm2, anti-Sld5 and anti-Psf3 antibodies, respectively. The corresponding region of the membrane was stained with Ponceau S as a loading control. Cellular DNA content was also measured (Supplementary Fig. 8c). E, 1% of the prepared extract; IP, 10% of the immunoprecipitate; S, 1% of the supernatant. **c**, Semi-conservative replication at ARS306 in YST575 (*JET1 GALp-sld2-11D GALp-DBF4*) cells. Cells pre-labelled with [^{13}C]glucose and [^{15}N]ammonium sulphate were arrested with α-factor with induction of Sld2-11D. (See Supplementary Fig. 8d for detailed experimental procedure.) **d**, Two-dimensional gel analysis of DNA replication intermediates at ARS305 and ARS306 in YST578 (*JET1 GALp-sld2-11D TetO$_2$-DBF4 GALp-SIC1ΔNT*) cells grown as in Fig. 3**c**.

Sld3 binds to Cdc45 (Fig. 2b; see also Supplementary Fig. 6) and GINS[26], and both Dpb11 and Sld2 bind to polymerase ε (Pol ε)[27,28]. Moreover, as we have shown here, Cdc45/Jet1 facilitates the interaction between Dpb11 and Sld3 (Supplementary Fig. 5b, c). However, neither enhanced Sld3 binding to Dpb11 upon addition of Jet1, nor increased affinity of Cdc45 to Sld3, Dpb11, GINS or Pol ε as a result of the *JET1* mutation, could be detected in an assay using purified proteins (our own unpublished data). We argue that these proteins form a fragile complex *in vivo* whose formation is promoted by phosphorylation-dependent interactions. *JET1* might increase the interaction between Sld3-2A and Dpb11 by stabilizing the multiprotein complex. We suggest that this complex formation is necessary for recruitment of replication machinery, such as Cdc45, GINS and Pol ε, to the putative helicase Mcm.

Cdc45 is well conserved from yeast to higher eukaryotes, and its association with chromatin is known as a landmark of CDK action in chromosomal DNA replication. By-pass of CDK requirement by the *JET1* mutation of *CDC45* suggests that CDK regulates the initiation step of DNA replication by controlling Cdc45 association with chromatin indirectly, even in higher eukaryotes. For example, TopBP1, a possible counterpart of Dpb11 in higher eukaryotes, has eight BRCT domains[29] and homologues of Sld2 and Sld3 may interact with TopBP1 to confer an involvement in DNA replication, as observed here.

METHODS

Strains and experimental procedures. The yeast strains used are listed in Supplementary Table 1. Conditions of cell growth, synchronization and other methods are described in Supplementary Methods.

Preparation of antibodies. The mouse monoclonal antibody specific to phospho-Thr 600 of Sld3 was raised against a mixture of CSEENVQVQA(pT)PAVKKR and SEENVQVQA(pT)PAVKKRTVTPNKKC. The rabbit polyclonal antibodies specific to phospho-Ser 622 of Sld3 were raised against QSIIE(pS)PLNFC, purified using phospho-peptide-conjugated columns and adsorbed to corresponding non-phosphorylated peptides. These phospho-specific antibodies were evaluated using Flag-tagged mutant Sld3. Rabbit anti-Sld3 antibodies were raised against GST–Sld3 produced in and purified from *Escherichia coli* and were purified using GST–Sld3 protein-conjugated columns.

Received 28 July; accepted 23 November 2006.
Published online 13 December 2006.

1. Schwob, E. & Labib, K. Regulating initiation events in yeasts. In *DNA Replication and Human Disease* 295–311 (Cold Spring Harbor Press, New York, 2006).
2. Kamimura, Y., Masumoto, H., Sugino, A. & Araki, H. Sld2, which interacts with Dpb11 in *Saccharomyces cerevisiae*, is required for chromosomal DNA replication. *Mol. Cell. Biol.* **18**, 6102–6109 (1998).
3. Masumoto, H., Muramatsu, S., Kamimura, Y. & Araki, H. S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature* **415**, 651–655 (2002).
4. Tak, Y. S., Tanaka, Y., Endo, S., Kamimura, Y. & Araki, H. A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. *EMBO J.* **25**, 1987–1996 (2006).
5. Kamimura, Y., Tak, Y. S., Sugino, A. & Araki, H. Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in *Saccharomyces cerevisiae*. *EMBO J.* **20**, 2097–2107 (2001).
6. Diffley, J. F. X. Regulation of early events in chromosome replication. *Curr. Biol.* **14**, R778–R786 (2004).
7. Blow, J. J. & Dutta, A. Preventing re-replication of chromosomal DNA. *Nature Rev. Mol. Cell Biol.* **6**, 476–486 (2005).
8. Sivaprasad, U., Dutta, A. & Bell, S. P. Assembly of pre-replication complexes. In *DNA Replication and Human Disease* 63–88 (Cold Spring Harbor Press, New York, 2006).
9. Bell, S. P. & Dutta, A. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* **71**, 333–374 (2002).

©2007 Nature Publishing Group

10. Kearsey, S. E. & Cotterill, S. Enigmatic variations: divergent modes of regulating eukaryotic DNA replication. *Mol. Cell* **12,** 1067–1075 (2003).
11. Mendez, J. & Stillman, B. Perpetuating the double helix: molecular machines at eukaryotic DNA replication origins. *Bioessays* **25,** 1158–1167 (2003).
12. Walter, J. C. & Araki, H. Activation of pre-replication complexes. In *DNA Replication and Human Disease* 89–104 (Cold Spring Harbor Press, New York, 2006).
13. Kanemaki, M. & Labib, K. Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO J.* **25,** 1753–1763 (2006).
14. Nguyen, V. Q., Co, C. & Li, J. J. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* **411,** 1068–1073 (2001).
15. Mendenhall, M. D. & Hodge, A. E. Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **62,** 1191–1243 (1998).
16. Araki, H., Leem, S. H., Phongdara, A. & Sugino, A. Dpb11, which interacts with DNA polymerase II (ε) in *Saccharomyces cerevisiae*, has a dual role in S-phase progression and at a cell cycle checkpoint. *Proc. Natl Acad. Sci. USA* **92,** 11791–11795 (1995).
17. Glover, J. N., Williams, R. S. & Lee, M. S. Interactions between BRCT repeats and phosphoproteins: tangled up in two. *Trends Biochem. Sci.* **29,** 579–585 (2004).
18. Zegerman, P. & Diffley, J. F. X. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* advance online publication doi:10.1038/nature05432 (13 December 2006).
19. Labib, K., Diffley, J. F. & Kearsey, S. E. G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nature Cell Biol.* **1,** 415–422 (1999).
20. Bishop, A. C. *et al.* A chemical switch for inhibitor-sensitive alleles of any protein kinase. *Nature* **407,** 395–401 (2000).
21. Masai, H. & Arai, K. Cdc7 kinase complex: a key regulator in the initiation of DNA replication. *J. Cell. Physiol.* **190,** 287–296 (2002).
22. Pacek, M., Tutter, A. V., Kubota, Y., Takisawa, H. & Walter, J. C. Localization of MCM2–7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol. Cell* **21,** 581–587 (2006).
23. Gambus, A. *et al.* GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nature Cell Biol.* **8,** 358–366 (2006).
24. Moyer, S. E., Lewis, P. W. & Botchan, M. R. Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl Acad. Sci. USA* **103,** 10236–10241 (2006).
25. Brewer, B. J. & Fangman, W. L. The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* **51,** 463–471 (1987).
26. Takayama, Y. *et al.* GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes Dev.* **17,** 1153–1165 (2003).
27. Masumoto, H., Sugino, A. & Araki, H. Dpb11 controls the association between DNA polymerases α and ε and the autonomously replicating sequence region of budding yeast. *Mol. Cell. Biol.* **20,** 2809–2817 (2000).
28. Edwards, S. *et al. Saccharomyces cerevisiae* DNA polymerase ε and polymerase σ interact physically and functionally, suggesting a role for polymerase ε in sister chromatid cohesion. *Mol. Cell. Biol.* **23,** 2733–2748 (2003).
29. Garcia, V., Furuya, K. & Carr, A. M. Identification and functional analysis of TopBP1 and its homologs. *DNA Repair (Amst.)* **4,** 1227–1239 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Diffley and P. Zegerman for sharing unpublished information; B. Stillman for anti-Orc6 antibody; A. Calzada, M. Foiani and M. Lopes for protocol of two-dimensional gel assay; H. Masukata for protocols of density labelling and detection of DNA in a CsCl density gradient; S. Endo, E. Nakashima, Y. Suzuki and S. Yamaguchi for technical assistance; Y. Tanaka for discussion; and J. Diffley, K. Furuya, H. Masukata and K. Sugimoto for critical reading of the manuscript. This study was partly supported by grants to H.A. and S.T. from Ministry of Education, Culture, Sports, Science and Technology, Japan.

Author Contributions S.M. isolated and characterized novel mutations of *DPB11*; Y.K. analysed the interaction of Sld3 with other proteins; K.H. analysed *in vitro* interaction between Dpb11 and Sld3; T.U. analysed the interaction between Sld3 and Dpb11, alanine-substituted *sld3* mutants and phosphorylation of Sld3; S.T. isolated and characterized the *JET1* mutation, and analysed the relationship between Cdc45/Jet1 and Sld3; H.A. organized this study and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.A. (hiaraki@lab.nig.ac.jp).

©2007 Nature Publishing Group

naturejobs

THE CAREERS MAGAZINE FOR SCIENTISTS

When academics talk casually about colleagues going into industry, they often cite salary as the prime motivation. But pay seems to be less important than other perks at some drug-development companies that made *Fortune* magazine's latest list of the top 100 employers in the United States. Genentech and Genzyme reported salaries of $76,000 to $86,000 for researchers. Yet despite being in the top 100 employers overall, none of these companies made it into the list of the 25 best-paid jobs.

So what is it about industry that makes it so attractive for academics? Time off, contact with management, and options to purchase stock, according to *Fortune* magazine's survey. Genentech, for instance, offers six weeks of paid sabbatical for every six years of service. Amgen provides 17 paid holidays — more than any other company in the top 100 — in addition to its standard three weeks of vacation per year. AstraZeneca holds 'town hall' meetings with employees and management, and employees have opportunities to breakfast with the US chief executive. Stock options, once a big draw to biotech companies, have played less of a role since the biotech bubble burst in 2001, but were still cited by Genzyme's employees as a key incentive.

Creating a good work environment has obvious benefits for a company, as it helps to draw top talent. But the competition is fierce for these coveted positions. In 2005, the year from which data were compiled, AstraZeneca received 631 applicants per job; Genentech received 135; and Genzyme, 120. The lowest of the four, Amgen, still had 31 applicants per job.

Companies in other industries that made the list paid even better — especially information-technology companies such as Adobe Systems, Yahoo and Microsoft, which all had salaries of more than $100,000. Engineers at EOG Resources also make well over $100,000 for their work — and get stock options and bonuses. The good news is that to capture the best talent — and bragging rights on the list — other companies might also start to offer better incentives.

Paul Smaglik, *Naturejobs* editor

Editor: Paul Smaglik
Assistant Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Business Development Manager:
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com
***Nature*events:**
Claudia Paulsen Young (+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com
France/Switzerland/Belgium:
Muriel Lestringuez (4994)

UK/Ireland/Italy/RoW:
Loredana Milanese (4944)
Nils Moeller (4953)
Scandinavia/Spain/Portugal:
Evelina Rubio-Morgan (4973)
Germany/Austria/The Netherlands:
Reya Silao (4970)
Online Job Postings:
Matthew Ward (+44 (0) 20 7014 4059)

European Satellite Office
Germany: Patrick Phelan
Tel: +49 89 54 90 57 11
Fax: +49 89 54 90 57 20
e-mail: p.phelan@nature.com

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

***Naturejobs* web development:** Tom Hancock
***Naturejobs* online production:**
Catherine Alexander

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe
e-mail: a.watanabe@natureasia.com

©2007 Nature Publishing Group

Don't stop me now

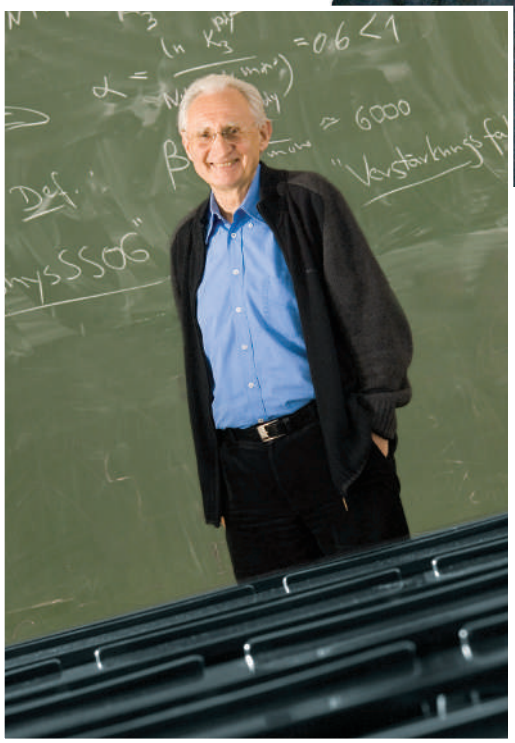
Unlike their US counterparts, European scientists can be forced to retire while they're still productive. But some in Germany are finding ways to go on. **Britta Danger** reports.

A year after the German physicist Theodor Hänsch won a Nobel prize he hit 65, and faced mandatory retirement from his posts at Ludwig Maximilians University in Munich and the Max Planck Institute of Quantum Optics in Garching. But Hänsch wanted to keep on working, putting the government in a potentially embarrassing position. The culprit was German civil-service regulations — because German university staff are classed as civil servants, they face mandatory retirement at 65. Numerous German researchers have been forced unwillingly into retirement as a result. But new initiatives could help German veterans continue to generate good science well into their golden years.

Hänsch himself has found a way to stay on in academia for at least five more years, but he owes his career extension to an exceptional initiative sponsored by the state government of Bavaria, the Carl Friedrich von Siemens Foundation and the Excellence Foundation for the Advancement of the Max Planck Society, which are collectively funding him to continue.

Moves are afoot to make such exceptions more the norm. To retain senior talent, the philanthropic Hertie Foundation, which funds neuroscience research, inaugurated Germany's first 'senior professorship' in July 2006. Only scientists over 60 were eligible to apply. Neurophysician Thomas Brandt, 63, a specialist in balance disorders, won the first award, which will allow him to keep working until at least the age of 68.

The Frankfurt-based foundation is currently the only German funding body that specifically gives grants to professors approaching retirement age. Without the grant, Brandt probably would have had to leave research, like dozens of experienced and highly qualified senior German researchers have to each year. Hänsch and Brandt are rare exceptions, but such cases may soon become more widespread. Private funding saved their research, and public agencies are getting increasingly interested as well.



D. ENDLICHER/AP

Nobel laureate Theodor Hänsch (top) and Jürgen Wolfrum: still at the peak of their powers.

Senior service

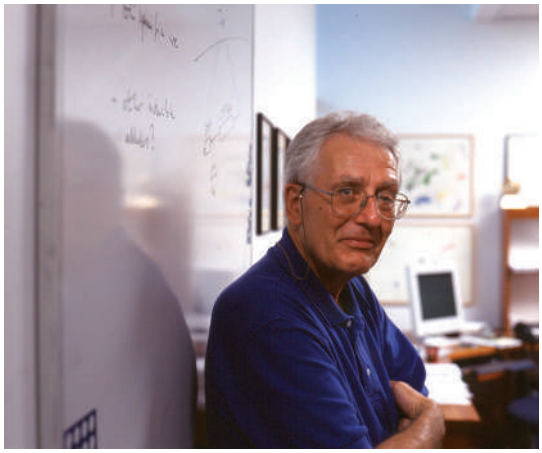
The situation is less clear outside Germany. Although a European directive has recently led member states of the European Union (EU) to implement age-discrimination laws, these tend to cover only employment rights up to the age of retirement. The result is somewhat confusing. In Italy, academics must either retire at 70, then request a two-year extension, or retire at 65 and continue for three years at full pay, without teaching. But the second option is rarely approved, as it places a higher pay burden on the universities. In Austria, academics must retire at 65, but can apply to extend their tenure to 68. In Switzerland, technical university professors must retire at 64 if they are women and 65 is they are men, although there have been some exceptions. The issue could even get more complex, as the shift towards an ageing population leaves countries considering whether to revamp their rules.

In Germany, a variation in the retirement rules among its 16 states could provide an impetus for change. In some states, professors can stay in office until a successor has been appointed. Some stick to the mandatory retirement laws; others allow an extension. In Bavaria, an extension until the age of 68 is possible, but is not guaranteed, regardless of how many publications or ongoing projects the researcher can show. And in Lower Saxony, the retirement age will become 68 this month.

Germany's science minister, Annette Schavan, says that Brandt's award could be pivotal for pushing back German scientists' retirement age. "He is a pioneer," she said in July 2005 at the ceremony marking the award of the first of the Hertie Foundation's senior-scientist grants, where she promised to change some of the outdated rules governing Germany's academic system. She was critical of the fact that staff are dismissed and projects interrupted just because a scientist is being forced to retire, saying the rules are out of date (see *Nature* **442,** 341; 2006). "The rules are from the last century and don't fit in our century any more," she said.

At the award ceremony, Schavan announced a programme through which senior professorships will be made available to outstanding older scientists throughout Germany. She noted that although the first

©2007 Nature Publishing Group



Klaus Rajewsky: never mind people's age — what research is getting done?

NEVER TOO LATE

Learning capability and efficiency decline as people get older — but they have many skills that can more than compensate, according to the late psychologist Paul Baltes, an expert in ageing and a leading voice in the German debate, who talked to *Nature* before his recent death. "Old people don't pick up new things so easily, he said. "But they have other strengths."

Baltes favoured a model that would encourage senior scientists to compete for a new position at a different university once they reach 65, to prove they can still be productive. He thought that the selection criteria should be stringent, and that only the best should actually make the cut. "Only a small proportion ages in excellence," he said. **B.D.**

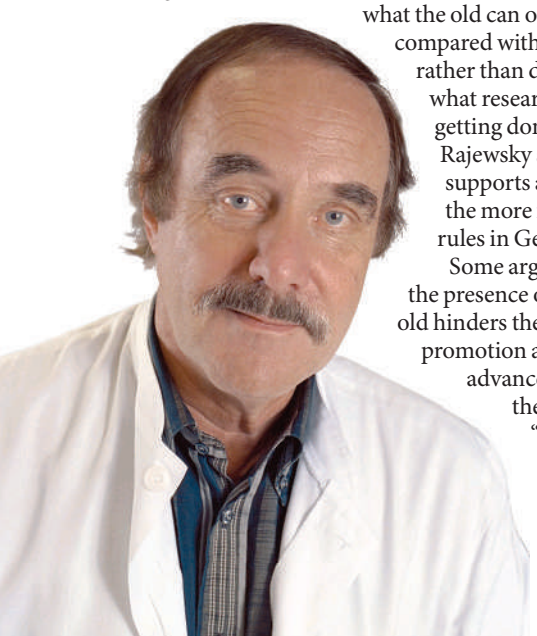
priority should be to improve career prospects for the young, she also plans to launch a programme later this year that would allow universities and research institutes to create senior professorships for top-class scientists over the age of 65.

Germany's current retirement rules risk causing a brain drain of eminent scientists from the country, which programmes to support older researchers could help to counter. Like many successful older German scientists, Hänsch had numerous offers from US universities to continue his career there; offers that increased after he won a Nobel prize. "It would have been extremely embarrassing to German science officials if I had left shortly after the award," he says.

Home from home

A mandatory retirement age was abolished in the United States in the 1990s following a lawsuit by researchers who claimed unfair discrimination. The country is now the most popular destination for older scientists with good publication records and an active network of international colleagues.

Geneticist Klaus Rajewsky, now aged 70, became famous for the knockout mice he developed while at the University of Cologne. In 2001, six months before he was due to retire, he moved to the Institute for Biomedical Research at Harvard Medical School in Boston, Massachusetts, taking with him 10,000 transgenic mice. "There is too much reflection about what the old can or cannot do compared with the young, rather than discussing what research is getting done," he says. Rajewsky strongly supports adopting the more flexible US rules in Germany.



Thomas Brandt: first to win a senior scientist's grant.

Some argue that the presence of the old hinders the promotion and advancement of the young. "To foster young scientists is the priority," says Beate Konze-Thomas, chief of the department of advancement of programmes and infrastructure of the DFG, Germany's main science-funding agency. The argument goes that if some senior scientists stay on for many years, they might prevent postdocs and young group leaders from becoming independent and showing their abilities to the full. But this does not usually apply to the best and most experienced older professors, says entomologist Bert Hölldobler of Arizona State University in Tempe. Most of them have spent much time and effort attracting and teaching as many talented young minds as possible, passing on their experience and the benefits of their reputation.

Experience counts

Hölldobler, 70, a native of Germany, was a professor of biology at Harvard University from 1973 to 1989, at which point he was appointed professor at the university of his student years in Wurzburg, Bavaria. He was amazed by its strict retirement rules. "Why did they make me come if they didn't want to exploit all of my capacities?" he asks. He left Wurzburg in 2004 to continue his scientific activities at Arizona.

"Older professors can really strengthen the standing of a university," says Brandt. He proposes a model for good coexistence of young and old: each university should create a few positions for excellent older scientists with designated laboratory space. This would require a small amount of money on top of the researchers' regular pension. The risk, says Brandt, would be very low, because the only scientists chosen would be those who are still successful in producing results and attracting grants. Continuing quality could be assured by regular evaluation (see 'Never too late').

There are other ways of getting round the age barrier and retaining the expertise of older scientists. The research institute Bioquant, created by the University of Heidelberg in 2001, specializes in interdisciplinary analysis of macromolecular complexes. Founding director Jürgen Wolfrum, 67, a specialist in laser applications, oversees research and leads the Bioquant's board of directors, along with two younger colleagues.

The institute will span not just disciplines, but ages. But the groups draw upon the experience of some scientists, such as Wolfram, who are past the traditional retirement age. "We create space for several junior groups," Wolfrum says. "The young scientists deliver new ideas, but I have a better general outline so I know about the practicability."

Wolfrum compares dedicated scientists to artists, who tend to carry on working regardless. "An enthusiastic scientist doesn't stop thinking just because he or she has reached the age of 65," he says. ■

Britta Danger was an intern in *Nature*'s Munich office.

©2007 Nature Publishing Group

The inside track from academia and industry

Great expectations

You know what you want from a job. But how do you convince an employer that your skills are relevant?

Jens-Peter Mayer

Working in the field of human resources has given me first-hand experience of the problems faced by both academics seeking jobs and employers seeking skilled workers. One of the most significant issues in the scientific sector is a mismatch in expectations. Often, a single vacancy in industry can generate hundreds of applications from academic scientists. Yet, despite this apparent wealth of choice for the prospective employer, none of the initial applicants will secure the job — leaving both sides disappointed.

This happens because industry often expects its candidates to have a set of specific attributes — it wants them to be highly motivated, to have a sound background in research, such as a good academic postdoc, and to have done at least one internship in industry. Meanwhile, what most academic candidates offer is a good research background, management experience and a host of 'soft' skills, such as writing and presenting data. But they often fail to explain how these skills will help them meet a company's needs. Sometimes they don't even realize that simply by virtue of their research they have gained highly relevant qualifications, such as a working knowledge of the requirements for 'good practice' in the drug industry.

"Academic skills can be transferable. Conducting research is equivalent to industrial work experience, and teaching reveals leadership skills."

As a human-resource professional, it is my job to try to balance the aspirations of academics aiming to leave university for a job in industry with the expectations of their prospective employers. To help bridge this gap, I have created a hypothetical candidate and a prospective employer, based on countless job hunts that my company has helped to facilitate.

THE CANDIDATE Jane has recently completed her PhD in chemistry. She has worked in a university laboratory, supervised students and supported her adviser. She is highly motivated and has a specialized knowledge of natural-product synthesis. As a result, she has several transferable skills. Conducting her own research is equivalent to industrial work experience, her teaching experience reveals that she has leadership skills, and managing a laboratory demonstrates that she can look after budgets.

But she doesn't realize that she needs to describe her skills in terms that make this clear to a prospective employer. So far, several job applications have failed to secure Jane a single interview, leaving her feeling dejected and questioning the value of her academic training — at least for the industrial market. So she continues to answer every industrial job advert she comes across, regardless of whether or not she meets the qualifications for the post on offer.

THE EMPLOYER But what of the prospective employers? In this hypothetical situation, let's consider just one, who advertised a position for someone to install and run a research laboratory to find target molecules for the development of a drug against cancer. The advert asked for several years' working experience under standardized research and production practices. The successful candidate would be responsible for strategic planning and management with regard to staffing and budget. The position demanded industry experience and management training.

But few, if any, applicants met the full profile. As a result, the employer was confused. Why did so many people send applications even though they did not meet the company's specific criteria? In fact, the whole experience has left the employer wondering whether the right candidate is out there at all.

Both the employer and the applicant had needs, but neither quite made the match. Why was this? Jane has failed, so far, because she does not really understand the demands of the industrial market and how to communicate her ability to fit into it. She feels helpless and frustrated and wonders whether her university training has perhaps failed her. The employer is frustrated, too. He does not know how to react to applications or how to cope with the hiring administrative process, and he fails to understand the candidate market. Both fail because they are not fully informed, and also because they feel pressed for time.

To connect the two communities — and to help real-life natural scientists and life-science companies overcome the barriers faced by Jane and her prospective employers — I will use this monthly column to create a platform for them. This will aim to help scientists from both academic and industrial backgrounds to improve their understanding of the market by providing tips on job search strategies, applications and industrial hiring trends.

I also aim to address employer issues — such as strategies for finding the ideal candidate and some alternate staffing approaches. Whatever the origin of the question, I aim to address possible solutions from both the employer's and jobseeker's perspectives. As a mediator, I aim to tackle hot topics in industrial staffing from both sides of the recruiter's door. I appreciate input, whether comments and questions or additional advice, at: naturejobs@kellyservices.de. ■

Jens-Peter Mayer is branch manager for Kelly Scientific Resources in Cologne, Germany.

©2007 Nature Publishing Group